

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 15:26:53 ; Search time 6842 Seconds  
(without alignments)  
11420.255 Million cell updates/sec

Title: US-10-019-633-1  
Perfect score: 1910  
Sequence: 1 ccacgcgtccgggtgtctcc.....aaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rtd.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1910	100.0	1910	6	AX081182	AX081182 Sequence
2	1836	96.1	2573	6	AX081194	AX081194 Sequence
3	1238	64.8	1988	6	AX081200	AX081200 Sequence
4	1101.4	57.7	2782	6	AX081196	AX081196 Sequence
5	1043.4	54.6	2520	6	AX357082	AX357082 Sequence
6	1041.8	54.5	2551	8	AY133852	AY133852 Arabidops
7	1041.8	54.5	2847	8	AY056301	AY056301 Arabidops
8	1041.8	54.5	2880	6	AX015734	AX015734 Sequence
9	1028	53.8	1803	6	AX507020	AX507020 Sequence
10	962.4	50.4	2482	6	AX081198	AX081198 Sequence
11	959	50.2	1816	6	AX081184	AX081184 Sequence
12	531.6	27.8	914	6	AX357098	AX357098 Sequence
13	498.6	26.1	42544	8	SPBC106	AL110295 S.pombe c
14	487.8	25.5	3667	10	BC049119	BC049119 Mus muscu
15	477.2	25.0	4761	8	SPADA1TCP	X91498 S.pombe ada
16	476.6	25.0	2732	5	BC044154	BC044154 Dario rer
17	475	24.9	71088	8	NCE7F18	AL389891 Neurospor
18	471.6	24.7	2857	9	HSU16270	U16270 Human AMP d
19	471.6	24.7	3359	9	S47833	S47833 Homo sapien
20	471.6	24.7	3386	9	HUMAMPD2	M91029 Human AMP d
21	471.6	24.7	3702	9	BC007711	BC007711 Homo sapi
22	470	24.6	3745	9	AX025706	AK025706 Homo sapi
23	465.4	24.4	34506	8	SC8337	Z46659 S.cerevisia
24	463.8	24.3	3354	8	YSCAMD	M30449 Yeast (S.ce
25	461.8	24.2	3237	3	AY051429	AY051429 Drosophil
26	458.8	24.0	2341	9	HUMAMPD1	M60092 Human myoad
27	448.4	23.5	2391	9	AK097077	AK097077 Homo sapi
28	447.8	23.4	2241	6	AX457110	AX457110 Sequence
29	441.2	23.1	2275	10	RATAMPDA	J02811 Rattus norv
30	439.6	23.0	4098	10	BC040366	BC040366 Mus muscu
31	438.8	23.0	2757	3	AF238311	AF238311 Dictyoste
32	436.2	22.8	3490	10	D85596	D85596 Mouse mRNa
33	422	22.1	3915	6	AX683214	AX683214 Sequence
34	422	22.1	3915	10	RNU90888	U90888 Rattus norv
35	416	21.8	3680	9	HUMAMPD3B	M84721 Human AMP d
36	416	21.8	3732	9	HUMAMPD3	D12775 Homo sapien
37	411.8	21.6	2945	10	BC007183	BC007183 Mus muscu
38	330.4	17.3	1928	10	BC016662	BC016662 Mus muscu
39	308.8	16.2	45189	3	U46753	U46753 Caenorhadi
40	305.4	16.0	2212	9	HUMAMPD3C	M84722 Human AMP d
41	256.4	13.4	35870	3	AC018749	AC018749 Leishmani
42	256.4	13.4	40020	3	AC068783	AC068783 Leishmani
43	234	12.3	44319	2	AC013191	AC013191 Drosophil
44	234	12.3	92969	2	AC010844	AC010844 Drosophil
45	234	12.3	163958	3	AC010705	AC010705 Drosophil

# ALIGNMENTS

RESULT 1  
AX081182  
LOCUS AX081182  
DEFINITION Sequence 1 from Patent WO0109305.  
ACCESSION AX081182  
VERSION AX081182.1 GI:13170066  
KEYWORDS Zea mays  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1  
AUTHORS Caspar,T., Falco,S.C., Sakai,H., Weng,Z. and Hu,X.  
TITLE Purine metabolism genes in plants

RESULT 2  
AX081194

LOCUS AX081194 2573 bp DNA linear PAT 27-FEB-2001  
DEFINITION Sequence 13 from Patent WO0109305.  
ACCESSION AX081194  
VERSION AX081194.1 GI:13170072  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1.  
REFERENCE  
AUTHORS Caspar, T., Falco, S.C., Sakai, H., Weng, Z. and Hu, X.  
TITLE Purine metabolism genes in plants  
JOURNAL Patent: WO 0109305-A 13 08-FEB-2001;  
E.I. DU PONT DE NEMOURS AND COMPANY (US); PIONEER HI-BRED  
INTERNATIONAL, INC. (US)  
LOCATION/Qualifiers  
FEATURES  
source 1..2573  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4577"  
BASE COUNT 716 a 523 c 588 g 694 t 52 others  
ORIGIN  
Query Match. 96.1%; Score 1836; DB 6; Length 2573;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 12 GGTTCCTCATGGGAGAGGAGTCAATAATGACCCCTGTACTCCAAACCTTAACCCCAA 71  
DB GTTGCTCATGGGAGAGGAGTCAATAATGACCCCTGTACTCCAAACCTTAACCCCAA 797  
QY 72 CCCGTTCACTTATGTGCGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGTGATGGCGT 131  
DB CCCGTTCACTTATGTGCGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGTGATGGCGT 857  
QY 132 TATCCATGTTATGCGGATAAAGATTGACGAGAGCATTTATCTGTGGCTGATGCTAC 191  
DB TATCCATGTTATGCGGATAAAGATTGACGAGAGCATTTATCTGTGGCTGATGCTAC 917  
QY 192 AACCTTCTCACTGACTTGCATTATATCTCCGAGTAACGGCTGAGGAGACACAGAAG 251  
DB AACCTTCTCACTGACTTGCATTATATCTCCGAGTAACGGCTGAGGAGACACAGAAG 977  
QY 252 TGTCTGCCAATATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCATCTGATGTTAAA 311  
DB TGTCTGCCAATATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCATCTGATGTTAAA 1037  
QY 312 TCGGATAGGGAATTTCTTGCCAGAGAAGCTGCCCCACATCGTGATTTTACAATGTGAG 371  
DB TCGGATAGGGAATTTCTTGCCAGAGAAGCTGCCCCACATCGTGATTTTACAATGTGAG 1097  
QY 372 GAAGGTCACACTCATGTTTCATCTTACGATGATGAATCAAAACATCTGTTGAGGTT 431  
DB GAAGGTCACACTCATGTTTCATCTTACGATGATGAATCAAAACATCTGTTGAGGTT 1157  
QY 432 CATAAATCCAACTAAGAAAGAACCTGATGAGGTGCTCATTTTTCAGAGATGGTACTTA 491  
DB CATAAATCCAACTAAGAAAGAACCTGATGAGGTGCTCATTTTTCAGAGATGGTACTTA 1217  
QY 492 TATGACTTTAAAGAGGTTTTTGTAGAGCTTGAAGCTTAACTGGGTATGATCTGAATTTGA 551  
DB TATGACTTTAAAGAGGTTTTTGTAGAGCTTGAAGCTTAACTGGGTATGATCTGAATTTGA 1277  
QY 552 TTGCTAGATGCCATGACAGACAAAGACATCTTATCTGTTTGGCAATTCATCTAAA 611  
DB TTGCTAGATGCCATGACAGACAAAGACATCTTATCTGTTTGGCAATTCATCTAAA 1337  
QY 612 ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCCTCAAAAGATAATCTTAT 671  
DB ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCCTCAAAAGATAATCTTAT 1397  
QY 672 TCAAGGCCGTTTTCTTGTGATGTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAA 731

DB 1398 TCAAGGCCGTTTTCTTGTGATGTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAA 1457  
QY 732 ATATCAGATGCGAGAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCA 791  
DB 1458 ATATCAGATGCGAGAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCA 1517  
QY 792 ACTTGCAAGTTGGATAGTAGGAAACAATGAATTCACAGTGGAAATGTTGTCTGGCTGGTTCA 851  
DB 1518 ACTTGCAAGTTGGATAGTAGGAAACAATGAATTCACAGTGGAAATGTTGTCTGGCTGGTTCA 1577  
QY 852 GATTCCAGCTTATATATGTTACAGGAATGGGTATCGTTACATCATTTCCAAATCT 911  
DB 1578 GATTCCAGCTTATATATGTTACAGGAATGGGTATCGTTACATCATTTCCAAATCT 1637  
QY 912 TCTTGACAAACATTTTCTCTCTTTTGTAGGTTACTATTGATCCAGCTTTCACACCCACA 971  
DB 1638 TCTTGACAAACATTTTCTCTCTTTTGTAGGTTACTATTGATCCAGCTTTCACACCCACA 1697  
QY 972 GCTCCATGTTCTCTGAAGCAGGTTGTAGGTTGGACCTGTTGATGATGAAGTAAACC 1031  
DB 1698 GCTCCATGTTCTCTGAAGCAGGTTGTAGGTTGGACCTGTTGATGATGAAGTAAACC 1757  
QY 1032 AGAAGGCCGTCACAAAGACATGCCACACCTTGAAACAGTGGACCAATGTTTCAACCC 1091  
DB 1758 AGAAGGCCGTCACAAAGACATGCCACACCTTGAAACAGTGGACCAATGTTTCAACCC 1817  
QY 1092 TGCATTTTCAATTTATGCTACTACTGCTATGCTTAACTTATTCACCTTAAACAGCTGCG 1151  
DB 1818 TGCATTTTCAATTTATGCTACTACTGCTATGCTTAACTTATTCACCTTAAACAGCTGCG 1877  
QY 1152 TGAGTCAAAGGGAATGACCACATATCAAAATTCCTGCCACATGCTGGAGGCTGGAGATGT 1211  
DB 1878 TGAGTCAAAGGGAATGACCACATATCAAAATTCCTGCCACATGCTGGAGGCTGGAGATGT 1937  
QY 1212 TGATCATTGCGCAGGACATTTCTTCTGCTGCAACAATATCAGATGGAATTAATCTAAG 1271  
DB 1938 TGATCATTGCGCAGGACATTTCTTCTGCTGCAACAATATCAGATGGAATTAATCTAAG 1997  
QY 1272 GAAGTCTCTGCTGCTCAGTACTTGTACTATCTTGGTTCAGATGCTGCGGATGTCGCC 1331  
DB 1998 GAAGTCTCTGCTGCTCAGTACTTGTACTATCTTGGTTCAGATGCTGCGGATGTCGCC 2057  
QY 1332 ATTGAGCAACAACCTCTTATTTCTTGACTATCATCGCAACCTTTTCCACGTTCTTCCA 1391  
DB 2058 ATTGAGCAACAACCTCTTATTTCTTGACTATCATCGCAACCTTTTCCACGTTCTTCCA 2117  
QY 1392 ACGAGGCTGGAATGCTCATTTATCTACGGATGACCTTTTGCAAAATTCACCTGACAAAAGA 1451  
DB 2118 ACGAGGCTGGAATGCTCATTTATCTACGGATGACCTTTTGCAAAATTCACCTGACAAAAGA 2177  
QY 1452 ACCATTGTTGGAAGAATACAGCATTTGCTTCCTGCTGGAAGCTCAGTTCTTGTGATTT 1511  
DB 2178 ACCATTGTTGGAAGAATACAGCATTTGCTTCCTGCTGGAAGCTCAGTTCTTGTGATTT 2237  
QY 1512 ATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTTCACATGCTCTCAAGCGCA 1571  
DB 2238 ATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTTCACATGCTCTCAAGCGCA 2297  
QY 1572 CTGGAATTTGAAGAACTACTTTCAAAGAGGACCTGCTGGAAATGATATTCACAGAACCA 1631  
DB 2298 CTGGAATTTGAAGAACTACTTTCAAAGAGGACCTGCTGGAAATGATATTCACAGAACCA 2357  
QY 1632 TGTAACCCCATCAGGTTTCAATTTTAGAGATGATCTGGGAATGAAATGAACTAGT 1691  
DB 2358 TGTAACCCCATCAGGTTTCAATTTTAGAGATGATCTGGGAATGAAATGAACTAGT 2417  
QY 1692 GTACTCTGACAATCAGATCTTAAATACAGAGCTGGACCTGTAAAGATGTCAGCCTCG 1751  
DB 2418 GTACTCTGACAATCAGATCTTAAATACAGAGCTGGACCTGTAAAGATGTCAGCCTCG 2477  
QY 1752 TGTATACAGACGAGTTGCGTTGTAGTGTGATGCGGAATTTATATCTCATGTTTGTGATG 1811





Db 1700 GG 1701

RESULT 4  
AX081196  
LOCUS AX081196  
DEFINITION Sequence 15 from Patent WO0109305.  
ACCESSION AX081196  
VERSION AX081196.1 GI:13170073  
KEYWORDS  
SOURCE Oryza sativa  
ORGANISM Oryza sativa

REFERENCE  
1. Caspar, T., Falco, S.C., Sakai, H., Weng, Z. and Hu, X.  
Purine metabolism genes in plants  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoidae; Oryzeae; Oryza.

AUTHORS  
TITLE  
JOURNAL  
E.I. DU PONT DE NEMOURS AND COMPANY (US) ; PIONEER HI-BRED  
INTERNATIONAL, INC. (US)

FEATURES  
Location/Qualifiers  
1..2782  
/organism="Oryza sativa"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4530"

BASE COUNT 798 a 640 c 591 g 753 t

ORIGIN

Query Match 57.7%; Score 1101.4; DB 6; Length 2782;  
Best Local Similarity 78.5%; Pred. No. 3.7e-219;  
Matches 1318; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY 13 GTTGCTCCATGGAGAGAGCTCAATAGACCCCTGACTCCAAACCTAACCCAC 72  
DB 768 GTTGCTCCATGGAGAGAGAACTAATGATCTAGTCTAGTCTCCAAACCTAAC 827

QY 73 CGGTTCACCTATGTGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGATGCGGT 132  
DB 828 CTTTCTATTACAGCAGCAGACTAAACTGAACATCATTTTGAATGTTGATGTT 887

QY 133 ATCCATGTTTATCGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGATCTACA 192  
DB 888 ATTCATGATATACCCCAATAAGACGCTAAAGAAAGAACTATCTCTGTGCTGATCTACT 947

QY 193 ACCTTCTTACCTGACTTGCATATATATCTCCAGTAAACGGCTGCAGGAGACACAGACT 252  
DB 948 ACCTTTTCTACTGATGATGACTATATCTCTGTGTTGGCTGCTGGGATATTCGAAT 1007

QY 253 GTCTGCCATAATCGGTAAATCTTCTGAGCATAAGTTTAAATTCATCTGATGTTAAAT 312  
DB 1008 GTATGTTATAACGTTTAAATCTTCTAGAACAGAAATCAATCTTCAATTTGATGCTAAT 1067

QY 313 GCGGATAGGAATTTCTGCCAGAAAGACTGCCCAACATCGTGATTTTACAAATGTCAGG 372  
DB 1068 GCGGATAGGAATCTACTTCTCAGAAAGCTGCACCCCATCGGACTTCTTACAATGTCAGG 1127

QY 373 AAGGTTGACATCATGTTTCAATTCAGCATGCAATCAAAACATCTGTTGAGGTTTC 432  
DB 1128 AAGGTTGATCTATCTGTTTCACTCTGATGATGATCAAGATTCAGAGATTTTGAATTT 1187

QY 433 ATAAATCCAAATCAAGAAAGAACTGATGAGGTGGTCAATTTTCAGAGATGCTACTTAT 492  
DB 1188 ATCAAGTCCAAAGTTGAGGAAAGAACTGACGAGGTTGATTTTATGAGATGCTACTAT 1247

QY 493 ATGACTTTAAAGAGGTTTTCAGAGCTTGGACTTAACCTGAGTATGATCTGAATGTTGAT 552  
DB 1248 TTGACTCTTAAGAGGTTTTCAGAGGTTTGGACTTGGACTTGGCTTATGACCTCAATGTTGAT 1307

QY 553 TTGCTAGATGTCATGACGACAAAGACATTTTCACTGTTTTCAGAAATTCATCTAAAT 612  
DB 1308 CTTCTAGATGTCATGACGACAAAGTACATTTCCATGCTTTTGAAGTTTCAATTTGAAG 1367

QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAACAGATAATCTTATT 672  
DB 1368 TATAATCTTGTGGCCAAATCCCGGCTGAGGAGATCTTTCTTAAACAGGACAACTTATT 1427

QY 673 CAAGCCGCTTTTCTGTGAGTGTGACAAAGCAAGTTTCTCTGACCTTTCTGTGACAAA 732  
DB 1428 CAAGCCGATTTTCTGTGAAATTTGCAAAAGAGATTTTCTGTGATCTTTGAAAGCAAGTAAA 1487

QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGACCAA 792  
DB 1488 TATCAGATGGCTGAGTATAGAAATCTATCTATGGGAGAAAGAAAGTGGATCAG 1547

QY 793 CTTGCAAGTTGGATAGTGAACAAATGTCACAGTGGAAATGTTGCTGCTGGCTTCCAG 852  
DB 1548 ATGGCAAGCTGGATAGTGAATGAATGTAATGTCACAGCAGAGATTTGTTGGTTAATTCAG 1607

QY 853 ATTCACGCTTATATATATGTCACAGAAATGGGTATCGTTACATCTTCCAAATCTT 912  
DB 1608 ATTCCTCGGATATACAAATGATATACAGGAGATGGAAACAATCAATTTCTCCAGAACCTC 1667

QY 913 CTTGACAACTTTTCTGTTCTCTTTTGGAGGTTTACTATTGATCCAGCTTTCACACCCACAG 972  
DB 1668 CTTGACAAATTTTCTGCTCTTTTGGAGTAACTGTTGATCTGCTTTCATCTCTCAG 1727

QY 973 CTCCATGTTCTTCTGAAGCAGGTTTGGAGGTTGGACCTGTTGATGATGAAGTAAACCA 1032  
DB 1728 CTCCATGTTTCTTGGCAACAGCTCGTTGGGCTGGATTTAGTGGATGATGAAGCAACCA 1787

QY 1033 GAAAGGCTCAACAAAGCACATGCCACACTGACAGTGGACCAATGTGTTTCAACCT 1092  
DB 1788 GAGAGACCCCAACAAACACATGCTACACCTGAGCAATGGAATTAATGTTTCAATCCA 1847

QY 1093 GCATTTTCATATATGCTGCTACTGCTATGCTAACTTATTCACCTAAACAGCTGCTG 1152  
DB 1848 GCATATGATATATGTTGATCTATTTGATCTAACTTGTGATCTGCTTCACTCTCAG 1907

QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTT 1212  
DB 1908 GAGTCAAGGATGATGACAAACAACTCAAACTCGTCCACACTGTTGGGAGCTGGAGATTT 1967

QY 1213 GATCATCTGGCAGGACATTTCTCTCTGTGACAAACATATCATATGGAATTAATCTAAG 1272  
DB 1968 GATCATCTGTGTCAGCATTTCTTACTTCTCATATATTTGCTCAGCGGTTTAAATTTAAG 2027

QY 1273 AAGTCTCTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332  
DB 2028 AAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2087

QY 1333 TTGAGCAACAACTCTCTTATTTTCTGACTATCATCGCAACCCCTTTTCCAACTGTTCTTCAA 1392  
DB 2088 TTGAGCAACAACTCATTTGTTTATTTGATTTATCAACGAAACCTTTTCCCAACATTTTCTCTA 2147

QY 1393 CGAGTCTGAATGCTCTCATATCTACGATGACCTTTGCAAAATTCACCTGACAAAAGAA 1452  
DB 2148 AGAGCTCTTAAAGTTTCTCTATCAACCGATGACCTTTGCAAAATTCACCTGACAAAAGAA 2207

QY 1453 CCATTTGTTGGAAGAAATACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512  
DB 2208 CCTTTGTTGGAAGAAATATAGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2267

QY 1513 TGGCAAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGCGGCAC 1572  
DB 2268 TGTGAATTTGCTAGGAATTTCTGTGACAGTCTGTTTCTCTCATAGCTCAAGTCACAC 2327

QY 1573 TGGATTTGTAAGAACTACTCTCAAAAGAGGACCTGCTGGAATGATTTTCAAGAACCAAT 1632  
DB 2328 TGGATTTGGAAGAACTACTCAAAAGAGGCTCATGATGGAATGATTTTCAAGAACCAAT 2387

QY 1633 GTACGACATCAGGTTTCAATTTTATGAGATGATCTGGAGAAATGAATGAATGAATGAAT 1691  
DB 2388 GTTCTCTACATCAGGATTTGAATTCGACACTATTTTGGAAAGAAAGAAATGGAGCTAAT 2446

```
RESULT 5
AX357082
LOCUS AX357082 2520 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 1 from Patent WO0206319.
ACCESSION AX357082
VERSION AX357082.1 GI:18674263
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eucosids II; Brassicales; Brassicaceae; Arabidopsi.
1
REFERENCE
AUTHORS Schulz, A.; Streber, W.; Hanke, C.; Schmidt, F. and Schubel, A.
TITLE Nucleic acid molecule which codes for a plant amp deaminase
JOURNAL Patent: WO 0206319-A 1 24-JAN-2002;
Aventis CropScience GmbH (DE)
FEATURES
source
1..2520
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
1..2520
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD23276.1"
/db_xref="GI:18674264"
/translation="MEPNYQLALALFASFAVSVGPFMHFKALNLVLKERKEN
PGDBPQNPFLVRRSQVRKQVNDQYRSPALDPATPFDGGGGGGDTGNSGHVY
VDEIPFLRLHPTEGSRASVHGASIRKTFSGFVRPISPKSPVASASAFESVEDDD
DNLNTEGSDASYLOANGNEMPADNEEQISMAASSMIRSHSODLHGVPDPIAA
DILRKEPEOTFVRLNPLEVPTSDVEAVKCIQECLEURKRYVQETVAPWEKEVIS
DPSTKPNTEPAHVPGQSDHCFEMQDGVHVFANKADELFPVADATAPFDLHH
VLKJAGNIRILCHRLVLEQKFNHLMNADREFLAQKSPHRDFTYRNRVDTHV
HHSACNQKHLRFILKSLRKQDPDEVIIFRDGTYLTLREVFSLEDLTGDLNVDLDDV
HADSTFHRDYLKYNPCGGRSREIFLKQDNLITQGRFLGEITKQVFSDEASKYQ
MAEYRISIVGRKMSWDOLASWIVNDLYSENVWMLIQLPRLNYIKDMGIVTSQNI
LONIFLPEATVPDPSQLHVLKOVVGLVDDESKPERRPKMPTPAQWTFNAF
NPAFSVYVYCANLYLVNKLRESKGMWITILRPHSGEAGDIDLAAATFLTCHSIAG
INRKSPLVQYLYLQAIGLWSPNSNLSFLDYHRNPPFPFLKLNWLSLSDTDPLQ
IHLTKEPLVEEYSIAASVWKLSDCLDEIARNISVYQSGFHALKSHWIGKDYKRGPD
GNDIHKTNVPHIRVFRDRTIWKEMQVYLGKAVISDEVVP"
BASE COUNT 699 a 541 c 572 g 708 t
ORIGIN
Query Match 54.8%; Score 1043.4; DB 6; Length 2520;
Best Local Similarity 76.0%; Pred. No. 4.5e-207;
Matches 1287; Conservative 0; Mismatches 406; Indels 0; Gaps 0;
QY 13 GTTGCTCCATGGGAGAGAGGTGCTAAATGACCCCTGACTCCAAAACCTAACCCCAAC 72
DB 799 GTTGACACATGGGAAAAGAGTCATCTGTATCTCTAGTACTCCAAAGCTTAATACAGAG 858
QY 73 CGGTTCACATATGTGCTGAACCAAAAGTCAGAGCATGTTTTCAAAACCTGTGATGCGGTT 132
DB 859 CCATTTGACACATCTCTCAGGGAATCTGTATCATTTGTTGAGATGCAAGATGGGTT 918
QY 133 ATCCATGTTATCGGATTAAGATTGTACGGAGAGCATTTATCTCTGGCTGATGCTACA 192
DB 919 GTCCACGTGTTTGCAAAATAAGATGCAAAAAGAGATCTCTTCCCGGTAGCTGATGCCA 978
QY 193 ACCTCTTCACTGACTTGATATATATCTCCGAGTAACGGCTGCAGGGAACCAACAT 252
DB 979 GCGTTTTTCACTGACTTGATCAACGTAACAGTCAAAAGTCAATAGCTGCAGAAACATCCCGACT 1038
QY 253 GTCTGCATAATCGGTAAATCTTCTTGAGCATAAAGTTTAAATTCATCTGATGTTAAAT 312
DB 1039 TTGTGCCACCGTCACTAGTCTCTCTAGAACAGAAATTAATCTCCATTTGATGCTTAAT 1098
QY 313 GCGGATAGGAATTTCTTGCCCAAGAGATGCCCCACATCGTGATTTTACAATGTCAGG 372
DB 1099 GCGGATAAAGAAATTTCTTGCTCAAAAAGATGCACCATCGTGATTTTATAACGTTAGG 1158
```

```
QY 373 AAGGTTGACACTCATGTTTCATCAITCAGCATGCGATGAATCAAAAACATCTGTTGAGGTTTC 432
DB 1159 AAGGTCGACACTCATGTCATCAITCAGCTGCGATGAACACCAACACCTTTTAAGGTTT 1218
QY 433 ATAAATCAAACTAAGAAAGAACCTGATGAGGTGGTCATTTTCAGAGATGTTCTTAT 492
DB 1219 ATTAAGTCAAAAGCTCCGAAAGAACCCGATGAGGTAGTAAATATTCGAGATGGAACATAT 1278
QY 493 ATGACTTTAAAGAGGTTTTCGAGAGCTTGACACTTAACGGGTATGATCTGAATCTTAT 552
DB 1279 TTGACCTTTGAGAGAGGTTTTCGAGAGCTTGATCTGATGAGGTATGACCTGAACCTCGAC 1338
QY 553 TTGCTAGATGTCATGACAGAAAAGACATTTTCATCGTTTTCACAAAATTCATATCAAAA 612
DB 1339 CTTTGGATGTTTCATGACAGAAAAGTACCTTTTCATCGTTTTCATGATGATGATGATGAT 1398
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAAATTTTCTCAAAACAAGATATCTTAT 672
DB 1399 TATAACCTTTGTTGTTCAAGTAGGCTTAGGGAGATTTTCTTAAACAGGATAATCTCATC 1458
QY 673 CAAGCCGTTTTCGCTGAGTTGACAAAGCAAGTTTTCCTGACCTTTCTGCTAGCAAA 732
DB 1459 CAAGGTCGATTTCTTGGTGAGATAACAAAGCAAGTCTTCTCTGACCTTTGAAGCTAGTAAA 1518
QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGAAACACAGATGAATGGGACCAA 792
DB 1519 TATCAGATGGCTGATATACAGATATCTATATATGCGAGAAAATAGCGAGTGGGACCAA 1578
QY 793 CTTGCAAGTGTGATAGTGAACAAATGAATGTCAGTGGAAATGTTGTCTGGCTGGTTTCAG 852
DB 1579 CTCGCTAGTTGGATTTGTAACAAATGATCTATACAGTGAAGATGTTGTCTGGTTAAATTCAG 1638
QY 853 ATTCACACCTTATATAATGTGTACAGAAATGGGTATCGTTTACATCATCTCCAAATCTT 912
DB 1639 CTCACACCTTTGTACAAATTTTACAGGACATGGGTATTTGACATCGTTTCCAGAAATATC 1698
QY 913 CTTGACAAATTTTCGTTTCTCTTTTTCAGGTTTACTATTTGATCCAGCTTTCACACCCACAG 972
DB 1699 CTGGACATATATTTTCTCTCTCTGTTGAGCCACGGTAGATCTCTGATTTCCATCTCCTCAG 1758
QY 973 CTCCTATGTTCTCTGAACAGAGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACCA 1032
DB 1759 CTCCTATGTTTTCGAGCAGGTTGTTGATTTGATTTGTTGATGATGAAGAACCAACCT 1818
QY 1033 GAAAGGCTCCAAACAGCAGATGCCACATCCCTGACAGTGGACCAATGTTGTTCAACCT 1092
DB 1819 GAAAGACCTCCCAACAAACACATGCCCACTCCAGCTCAATGGACTTAACGATTTCAATCTCT 1878
QY 1093 GCATTTTCATATTTATGCGTACTTACTGCTATGCTAACTTATTCACCTTAAACAAAGCTGCGT 1152
DB 1879 GCATTTTCGATTTATGTTCTACTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1938
QY 1153 GAGTCAAAAGGGAATGACCACTATCAAAATTCGCTTCCATGTCAGTGGAGAGGCTGGAGATGTT 1212
DB 1939 GAGTCAAAAGGCGATGACTACTATACGCTACGACACATCTCTGGAGAGGCTGGTGCATTT 1998
QY 1213 GATCACTTTGGCAGCAGATTTCTTCTCTGTGCAACATATCACATGGGAATTAATCTAAGG 1272
DB 1999 GACCACTTTGGCTGCTACGTTTCTTAACTGCGCATAGCATGGAATCAATCTGCGA 2058
QY 1273 AAGTCTCTGCTGCTTTCAGTACTTGTACTTATCTTGTGTCAGATTTGTTCTGGCGATGTCGCCA 1332
DB 2059 AAGTCTCTGCTGCTTTCAGTACTTGTACTTACCTCGCCAGATTTGTTCTGGCCATGTCACCA 2118
QY 1333 TTGAGCAACAACTCTCTTATTTCTTCAATATCAATCAACACCTTTTCCAAAGTTCCTTCCAA 1392
DB 2119 CTGAGCAACAACTCTCTTGTCTTAGATTAACACCGGAACCCGTTTCTCTGTTGTTTCTTCT 2178
QY 1393 CGAGGCTGGAATGCTCTCATTTCTAGGATGACCTTTTGCAAAATTCACCTGCAACAAAGAA 1452
DB 2179 AGAGGCTCAATGTTTCTCTGCTACTGATGACCCCTTTCAGATTCACCTTAACCTAAGAA 2238
```



QY 433 ATAAATCCAACTAAGAAAAAGAACTGATGAGGTGGTCAATTTTCAGAGATGTTACTTAT 492  
DB 1219 ATTAAGTCAAGCTCGGAAAGAACCGATGAGGTGTAATATTCGAGATGGAACATAT 1278  
QY 493 ATGACTTTTAAAGAGGTTTTGAGAGCTTTGGACTTAATCTGGGTGATGATCTGAATGTGAT 552  
DB 1279 TTGACCTTGAGAGAAAGTTTTGAGAGCTTTGGATCTGACTGGATATGACCTGGAAGCTGCAC 1338  
QY 553 TTGCTAGATGCTCCATCCAGACAAAGACATTTTCATCGTTTTGACAAATTTCAATCTAAAA 612  
DB 1339 CTTTGGATGTTTCATCGACACAAAGATACCTTTTCATCGTTTTGATAGTTTCAACCTAAAG 1398  
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAAACAGATAATCTTAT 672  
DB 1399 TATACCTTTGGTCAAGTAGGCTTAGGAGATTTTCTTAACAGGATATCTCATC 1458  
QY 673 CAAAGCCGTTTTCTCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTCTAGCAAA 732  
DB 1459 CAAAGTCGATTTCTTGGTGAGATAACAAAGCAAGCTTCTCTGACCTTGAAGCTAGTAAA 1518  
QY 733 TATCAGATGCGAGATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCA 792  
DB 1519 TATCAGATGCGTGAATACAGAAATATCTATATATGCGAGAAAAATGAGCGAGTGGGACCA 1578  
QY 793 CTTGCAAGTTGGATAGTGAACAAATGAATTCAGAGTGGAAATCTTCTCTGGCTGGTTCA 852  
DB 1579 CTCGCTAGTTGGATTTGGAACATGATCTATACAGTGAAGATTTGCTGTGGTTAATTCAG 1638  
QY 853 ATTCCACGTTTATATATGTGTACAAGGAAATGGGTATCGTTACATCATTTCCAAATCTT 912  
DB 1639 CTCACACGCTTGACAAATTTACAAGGACATGGGTATTGTGACATCGTTCCAGATAATC 1698  
QY 913 CTTGACAAATTTCTCTCTCTTTTGGAGTTTACTTATGATCCAGCTTTCACACCCACAG 972  
DB 1699 CTGGACAATATATTTCTCTCTTTTGAAGCACGCTAGATCTCTGATTTCCATCCTCAG 1758  
QY 973 CTCATGTTCTCTGAAGCAGGTTGTAGGTTGGACCTGTTGATGATCAATCAAAAGTAAACA 1032  
DB 1759 CTCATGTTTTTGAAGCAGGTTGTGGATTTGATTTGGTGTGATGATGAAGCAACCT 1818  
QY 1033 GAAAGCGCTCCAAACAAAGACATGCCCAACCTGAAACAGTGAACCAATGTTTCAACCT 1092  
DB 1819 GAAAGCGTCCCAACAAACACATGCCCACTCCAGCTCAATGGACTAAGCAATCAATCCT 1878  
QY 1093 GCATTTTTCATATATGCTGCTACTGCTATGCTTAATCTTACCTTAACAGCTGCGT 1152  
DB 1879 GCATTTTCTGTTATGCTACTATTTGTTATGCTTAACCTCTATGTTTAAATAGCTTCA 1938  
QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTT 1212  
DB 1939 GAGTCAAGGGAATGACTACTATCACTGCTACGACCACTTCTGGAGAGGCTGGTGAAT 1998  
QY 1213 GATCACTTGGCAGCAGATTTCTTCTCTGCAACATATCAATGAAATTAATCTAAG 1272  
DB 1999 GACCACTTGGCTGCTACGTTTCTTAACATGCCATAGCATCGCATGCAATCAATCTGCGA 2058  
QY 1273 AGTCTCTGCTGCTCAGTACTGTTACTATCTTGGTTCAGATGCTGGGATGTCCTCA 1332  
DB 2059 AAGTCTCTGCTGCTCAGTACTGTTACTACCTCGCCAGATGCTGGCCATGTCACCA 2118  
QY 1333 TTGAGCAACAACTCTTATTTCTGACTATCATCGCAACCTTTTCCAAAGCTTTCTCCAA 1392  
DB 2119 CTGAGCAACAACTCTTGTGTTCTAGATTACCCAGGAAACCGTTCTGTTGTTTCTTA 2178  
QY 1393 CGAGGTCTGAATGCTCATATCTACGAGTACCCCTTTGCAATTCACCTGACAAAGAA 1452  
DB 2179 AGAGGTCTCAATGTTTCTGCTACTGATGACCCCTTTCAGATTCACTTAACAAAGAA 2238  
QY 1453 CCATTGGTGAAGATACAGCATTTGCTGCTGCTGGAAGCTGATTTCTGTGATTTA 1512  
DB 2239 CCTCTCGTGAAGAGATAGCATAGTGTGATCAGTTTGAAGCTGATGCGTGTGACCTG 2298

QY 1513 TCGCAATTTGGAGGAACCTGTTTACCAATCTGGGTTTTTACATGCTCTCAAGCGCAC 1572  
DB 2299 TGGGAGATAGCTCGTAACTCAGTGTACAGTCAGGTTTTCTCACAGCCTGAAGTCGCAC 2358  
QY 1573 TGGATTGTTAAGAACTACTTCAAAAGAGGAGCTGCTGGAATGATATTTACAGAACCAAT 1632  
DB 2359 TGGATTGTTAAGAAATTTACTACAAAAGAGGAGCTGATGGAACGACATTTCAAAAACAAAC 2418  
QY 1633 TACCGCACATCAGGTTTCAATTTAGAGAGATGATCTGAGAAATGAATGAACCTAGTG 1692  
DB 2419 GTGCCACACATAAGGTTGGAGTTCGTGACAGATCTGAAAAGAGGAGATGCAACAGTT 2478  
QY 1693 TACTCTGACAATG 1705  
DB 2479 TATCTGGGCAAGG 2491

RESULT 7  
AY056301  
LOCUS  
DEFINITION  
Arabidopsis thaliana putative AMP deaminase (At2g38280) mRNA.  
complete cds.  
ACCESSION  
VERSION  
AY056301.1 GI:15810524  
KEYWORDS  
SOURCE  
FLI CDNA.  
ORGANISM  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi  
1 (bases 1 to 2847)  
REFERENCE  
AUTHORS  
Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,  
Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L.,  
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,  
Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,  
Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,  
Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,  
Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Arabidopsis Full Length cDNA Clones  
Unpublished  
2 (bases 1 to 2847)  
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
Riken Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

TITLE  
JOURNAL  
COMMENT  
Direct Submission  
Submitted (12-SEP-2001) Plant Gene Expression Center, 800 Buchanan  
Street, Albany, CA 94710, USA  
Riken Genomic Sciences Center (GSC) members carried out the  
collection and clustering of RAFL cDNAs (RAFL CDNA : 'RIKEN  
Arabidopsis Full-Length cDNA') : Seki, M., Narusaka, M., Ishida, J.,  
Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,  
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the  
sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J.,  
Banno, P., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S.,  
Quach, H.L., Tang, C.C., Toriumi, M., Yamamura, Y., Yu, G., Yu, S.,  
Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C.,  
Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M.,  
Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W.,  
Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to  
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)  
contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis  
genome submitted to GenBank.

## FEATURES

Source  
Location/Qualifiers  
1..2847  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosome="2"  
/clone="RAF107-08-G24 (R10978)"  
/note="this clone is in a modified pBluescript vector (FLC-1) as a BamHI/XhoI insert."  
ecotype: Columbia  
1..2847  
/gene="At2g38280"  
1..83  
/gene="At2g38280"  
84..2603  
/gene="At2g38280"  
/codon\_start=1  
/product="putative AMP deaminase"  
/evidence=experimental  
/protein\_id="AA107150.1"  
/db\_xref="GI:15810525"  
/translation="MEPNIQALALFCAAFVAVSGFFMHFKALNVLVLRGKERKEN  
PDGEPQNTLVRRRSQVRKRVNDQYGRSPASLPDPTDGGGGGDDTGRSNGHVV  
VDEIPEGLPLHTRSEGRASVHGASSIRKTSFVRPIIPKSPVASASAFESVEESDD  
DNLNTEGLDASYLAQNDNEMPADANEQISMAASSMIRSHSVSGDLHGVPDPIAA  
DILKEPQETVRVNLVEVTSDEVEAYKCLQECLELRKRYFEIVAPWEKEVIS  
DPSTPKPNTPEAHYPOQKSDHCFEMQGVVHVFANKDAKEDLPVADATAFETDLHH  
VHSAQNKIRTLCHRRLLFLKRLKPEBWIIFRDTYTLTLREVPSLDTLQVFNVRKVDTH  
HLKAMQNKHLRFTKSLRKEPDRLEIFLKDNLIOGRFLEITKQVPSDLASKYO  
HAEKSTFRFEDKFNLYKPCGOSRLREIFLKDNLIOGRFLEITKQVPSDLASKYO  
MAEYISIIYGRKMSWDQLASWVNNDLIYSENVWLIQLPRLYNIYKMGIVTSQNI  
LDNIFLPEIDPDSPHQLHVFLLQVVDLVDDESKPERPKHMPKTPAQNWNAF  
NPAFSYVYCYANLYVNLKRESGMTITLPHSGAGDIDHUAATFLTCHSTAHG  
INLRSPVLQVLYLAQGLAMSPLSNLSFLDHYRNPFPVFFLGLNVLSTDDPLQ  
IHLTKPELVEEYSIAASVWKLSDCLCEIARNVSVYSGFSHALKSHWIKGDIYKRGPD  
GNDIHKTVPHIRVEFRDTINKEEMQVYLGKAVISDEVVP"  
2604..2847  
3'UTR  
/gene="At2g38280"  
BASE COUNT 817 a 593 c 641 g 796 t  
ORIGIN

Query Match 54.5%; Score 1041.8; DB 8; Length 2847;  
Best Local Similarity 76.0%; Pred. No. 9.6e-207;  
Matches 1286; Conservative 0; Mismatches 407; Indels 0; Gaps 0;  
QY 13 GTTGTCTCATGGGAGGAGGTCTAAATGACCCCTGTACTCCAAAACCTAAACCCCAAC 72  
DB 882 GTTGACCATGGGAAAAGAGTCATCTGTATCTGTACTCTCAAGGCTTAATACAGAG 941  
QY 73 CCGTTTCATTATGTGCTGAAACCAAGTCAGAGCATGTTTTCCAAACTGTTGATGCGGTT 132  
DB 942 CCATTGACACATCTCTCAGGAAAACTCTGATCATTTGTTTGAGATGCAAGATGGGTT 1001  
QY 133 ATCCATGTTTATGCGGATAAGATGTTACGGAGAGCATTTATCTCTGTGGCTGATGCTACA 192  
DB 1002 GTCCACGTTGTTGCAATAAAGATGCAAAAGAGATCTCTCCCGGTAGCTGATGCCACA 1061  
QY 193 ACCTCTTCACGTACTGCTGATATATCTCCGAGTAAACGGCTGAGGGAACAACAAGAACT 252  
DB 1062 GGTGTTTTCACGTACTGCTGATCACTCAAGTCATAGCTGACGAGAAACATCCGGACT 1121  
QY 253 GTCTGCCATAATPCGGTTAAATCTCTTGAGCAATAAGTTTAAATTCATCTGATGTTAAAT 312  
DB 1122 TTGTGCCACCGTCGACTAGTGTCTCTAGAACACAGAAATTTAATCTCCAATTTGATGCTTAAT 1181  
QY 313 GCGGATAGGAAATTTCTGCCACAGAGACTGCCACACATCGGTATTTTCAATGTGACGG 372  
DB 1182 GCGGATAAAGATTTCTTGCTCAAAAAGTGCACACATCGGTGATTTTATAACGCTTAGG 1241  
QY 373 AAGGTTGACACTCATGTTTCATCATTCAGCATGCAATGAAATCAAAACATCTGTTAGGTTTC 432  
DB 1242 AAGTCGACACTCATGTGTCATTCATTCAGCTTGCATGGAACCAACACACCTTTTAAGGTTT 1301

QY 433 ATAAATCCAAACTAAGAAAGAACCTGATGAGGTGGTCAATTTTCAGAGATGGTACTTAT 492  
DB 1302 ATTAAGTCAAAAGCTCCGAAAAGAACCCGATGAGGTGTAAATATTCGAGATGGAACATAT 1361  
QY 493 ATGACTTTAAAGAGGTTTTTGGAGAGCTTGACATTAACCTGGGTATGATCTGAATGTTGAT 552  
DB 1362 TTGACCTTGAGAGAAGTTTTTGGAGAGCTTGATCTGACTGGATATGACTGACCTGACCTGCAC 1421  
QY 553 TTGCTAGATGTCATGACAGACAAAGACATTTTCATCGTTTTTGACAAATTCATTAATAAAA 612  
DB 1422 CTTTGGGATGTTTCATGACAGACAAAGATACCTTTTCATCGTTTTGATAAGTTTCAACCTAAAG 1481  
QY 613 TACATCCATGTCGCCAAAGTAGGCTCAGAGAAATTTTCTCAACACAGATAATCTTATT 672  
DB 1482 TATAACCCCTTGTGGTCAAAAGTAGGCTTAGGAGATTTTTCTTAAACAGGATAATCTCATC 1541  
QY 673 CAAGGCCCTTTCTGTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTTCTGCTAGCAAA 732  
DB 1542 CAAGGTCGATTTCTTGGTGAGATAACAAGCAAGTCTTCTCTGACCTTTGAAGCTAGTAAA 1601  
QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCAA 792  
DB 1602 TATCAGATGGCTGAATACAGAAATATCTATATATGACAGAAAAATGAGCGAGTGGACCAA 1661  
QY 793 CTTGCAAGTTGGATAGTGAACAAATGAATGACAGTGGAAATGTTGTCTGGCTGGTTGAG 852  
DB 1662 CTGCGTAGTTGGATTTGTGAACAATGATCTATACAGTGAAGATGTTGTCTGGTTAAATTCAG 1721  
QY 853 ATTCACGCTTATATAATGTGTACAGAAATGGGTATCGTTTACATCATTCACAAATCTT 912  
DB 1722 CTCCACGCTTGTACAACTTACAGGACATGGGTATTTGACATCATCGTTTCCAGAAATATC 1781  
QY 913 CTTGACAACTTTTGGTTTCTCTCTTTTGGAGGTGTACTATTGATCCAGCTTTCACACCCACAG 972  
DB 1782 CTGACACATATATTCATTTCTCTCTGTTTGAAGCCAGGTAGATCTCTGATTTCCATCTCAG 1841  
QY 973 CTCCATGTTCTTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAAGTAAACCA 1032  
DB 1842 CTCCATGTTTTTTTGAAGCAGGTTGTGGATTTGATTTGGTGTGATGAAAGCAAACT 1901  
QY 1033 GAAAGGCTCCAAACAAAGCACATGCCACACCTGAAACAGTGGAGCAATGTGTCAACCT 1092  
DB 1902 GAAAGACCTCCCAACAAACACATGCCCACTCCAGCTCAATGGACTTAACGCAATCAATCT 1961  
QY 1093 GCATTTTCATATATGCGTACTTCTGCTATGCTAACTTATTCACCTTAAACAAAGCTCGT 1152  
DB 1962 GCATTTTCGATATATGCTACTATTTGTTATGCTAACTTATGTTTAAATAGCTTGA 2021  
QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGCTCCATGTCGATGAGAGGCTGGAGATGTT 1212  
DB 2022 GAGTCAAGGCGCATGACTACTATCACTAGGACCACTTCTGGAGAGGCTGGTGACATT 2081  
QY 1213 GATCACTGGCAGCGACATTTCTTCTGTCACACATATCACATGGAATTAATCTAAGG 1272  
DB 2082 GACCACTGGCTGTGCTAGCTTTCTAACTGATGATGATGATGATGATGATGATGATGATGAT 2141  
QY 1273 AAGTCTCTGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1332  
DB 2142 AAGTCTCTGCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2201  
QY 1333 TTGAGCAACAACTCTTATTTTCTGATCATCATCGCAACCCCTTTTCCAACTGTTTCCAA 1392  
DB 2202 CTGAGCAACAACTCTTGTGTTCTAGATTACCAACCGAACCCGTTTCTGCTGCTGCTGCTGCT 2261  
QY 1393 CGAGGTCTGAATGCTCATTTATCTACGATGACCCCTTTGCAAAATTCACCTGACAAAGAA 1452  
DB 2262 AGAGGTCTCAATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2321  
QY 1453 CCATTTGTTGGAAGATACAGCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1512  
DB 2322 CCTCTGTTGGAAGAGTATAGCATAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2381  
QY 1513 TGCAGAAATTGCGAGGAACTCTGTTTACCAATCTGGGTTTTTTCATGCTCTCTCAAGGCGCAC 1572



```
Db 2076 GACCACTTGGCTGCTACGTTTCTTAACATGCCATAGCATCGCATGGAATCAATTCGCA 2135
Qy 1273 AAGTCTCTGCTGCTCAGTACTTGTACTATCTTGTGTCAGATGGTCTGGCGATGTCCTCA 1332
Db 2136 AAGTCTCTGCTGCTCAGTACTTGTACTATCTTGTGTCAGATGGTCTGGCGATGTCCTCA 2195
Qy 1333 TTGAGCAACAACCTCTTATTTCTTGTACTATCTTGTGTCAGATGGTCTGGCGATGTCCTCA 1392
Db 2196 CTGAGCAACAACCTCTTATTTCTTGTACTATCTTGTGTCAGATGGTCTGGCGATGTCCTCA 2255
Qy 1393 CGAGGTCTCAATGCTCTATTAATCTACGATGACCTTTTCAAAATTCACCTGACAAAGAA 1452
Db 2256 AGAGGTCTCAATGCTCTTGTCTACTGATGACCTTTTCAAAATTCACCTGACAAAGAA 2315
Qy 1453 CCATTGGTGAAGAAATACAGCAATGCTGCTTGTGTCAGATGGTCTGGCGATGTCCTCA 1512
Db 2316 CCTCTCGTGAAGAGTATAGCATAGCTGATCAGTTTGAAGCTGAGTGGCTGACCTG 2375
Qy 1513 TGGCAATTTGCGAGGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGGCAC 1572
Db 2376 TGGCAATTTGCGAGGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGGCAC 2435
Qy 1573 TGGATTGGTAAGAACTACTTTCAAAAGAGGACCTCTGGAATGATATTCACAGAAACCAAT 1632
Db 2436 TGGATTGGTAAGAACTACTTTCAAAAGAGGACCTCTGGAATGATATTCACAGAAACCAAT 2495
Qy 1633 GTACCGCACATCAGGGTTCAAATTTAGAGAGATGATCTGGAGAAATGAAATGAAATCTAGTG 1692
Db 2496 GTGCCACACATAAGGCTGAGTTCCTGTACACAGTCCGTGGAAGAGGAGATGCAACAGTT 2555
Qy 1693 TACTCTGACAAATG 1705
Db 2556 TATCTGGGCAAGG 2568
```

```
RESULT 9
AX507020
LOCUS AX507020 1803 bp DNA linear PAT 27-SEP-2002
DEFINITION Sequence 1715 from Patent WO0216655.
ACCESSION AX507020
VERSION AX507020.1 GI:23388257
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE Harper, J.F., Kreps, J., Wang, X. and Zhu, T.
AUTHORS Stress-regulated genes of plants, transgenic plants containing
TITLE same, and methods of use
JOURNAL Patent: WO 0216655-A 1715 28-FEB-2002;
The Scripps Research Institute (US); Syngenta Participations AG
(CH)
FEATURES
source Location/Qualifiers
1..1803
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
BASE COUNT 524 a 384 c 378 g 517 t
ORIGIN
Query Match 53.8%; Score 1028; DB 6; Length 1803;
Best Local Similarity 76.2%; Pred. No. 7.3e-204;
Matches 1265; Conservative 0; Mismatches 395; Indels 0; Gaps 0;
Qy 13 GTTGTCTCATGGGAGAGGATCAATAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 72
Db 100 GTTGACCATCGGAAAGAAAGATCATATCTGATCTCTAGTACTCCAAAGCCTAATACAGAG 159
Qy 73 CGGTTCACTTATGTGCTTGAACCAAGTCAGAGCATGTTTCCAACTGTTGATCGCGTT 132
```

```
Db 160 CCAATTGCACACTATCTCTCAGGAAATCTGATCAATGTTTGTGAGATCCAAAGATGGGTT 219
Qy 133 ATCCATGTTTATGGGATAAAAGATTGTACGGAGAGCAATTTATCTGTGGCTGATGCTACA 192
Db 220 GTCCACGTTGTTGCAAAATAAAGATGCAAAAAGAGATCTCTTCCCGGTAGCTGATGCCACA 279
Qy 193 ACCTTCTTCACTGACTTGCATTTATTTCTCCGAGTAACGGCTGCAGGGAACACAAGAACT 252
Db 280 CGGTTTTTCACTGACTTGCATCAGCTACTCAAGTCTAGCTGCAGGAAACATCCGGACT 339
Qy 253 GTCTGCCATAATCCGTTTAAATCTTCTTGAGCATTAAGTTTAAATTCATCTCATGTTAAAT 312
Db 340 TTGTGCCACCTGCACTAGTGTCTCTAGAACAGAAATTTATCTCCATTTGATGCTTAAT 399
Qy 313 CGGATAGGGAATTTTGTGCCAGAACTGCCCAATCGTGATTTTAAATCAATGTTCAGG 372
Db 400 GCGGATAAGAAATTTCTTGCTCAAAAAGTGCACCAATCGTGATTTTATAACGTTAGG 459
Qy 373 AAGTTGACACTCATGTTTCATCATTTCAGCATGCAATGCAATCAAAAACATCTCTTGAGGTT 432
Db 460 AAGTTCGACACTCATGTTTCATCATTTCAGCTTGCATGAACCAAGAAACACCTTTAAGGTT 519
Qy 433 ATAAATCCAAACTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 492
Db 520 ATTAAGTCAAAAGCTCCCGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 579
Qy 493 ATGACTTTTAAAGAGGTTTGTGAGAGCTTGACCTTAACCTGGGTATGATCTCAATGTTGAT 552
Db 580 TTGACCTTTGAGAGAAATTTTGTGAGAGCTGGATCTGACTGGATATGACCTGAACGTCAC 639
Qy 553 TTGCTAGATGTCATGACAGCAAAAGCAATTTTCATCGTTTGTGACAAATTCATCTCAAAA 612
Db 640 CTTTGTGATGTTTCATGACAGCAAAAGTACCTTTTCATCGTTTGTGATAAGTTCAACCTAAG 699
Qy 613 TACAATCCATGTCGCCAAAGTAGGCTCAGAGAAATTTTCTCAAGCAAGATAATCTTATT 672
Db 700 TATAACCTTGTGTTCAAGTAGGCTTAGGAGATTTTCTTAAACAGGATAATCTCATC 759
Qy 673 CAAGCCGTTTCTTGTGCTGATTTGACAAAGCAAGTTTTTCTGACCTTTCTGTAGCAAAA 732
Db 760 CAAGGTCGATTTCTTGTGAGATAAAGCAAGCAAGTTCTCTGACCTTTGAAGCTAGTAAA 819
Qy 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGAAAGAAACAGAGTGAATGGACCAA 792
Db 820 TATCAGATGGCTGAATACAGAAATCTATATATATGCGAGAAAAATCAGCGAGTGGACCAA 879
Qy 793 CTTCAAGTTGAGTAGTGAACAATGCAATGCAAGTGAATGTTGTTGCTGGCTGCTCAG 852
Db 880 CTCGCTAGTTGGATTGGAACAATGATCTATACAGTGAAGATGTTGTTGTTTAAATTCAG 939
Qy 853 ATCCACGCTTATATAATGTTGTAAGAAATGGGTATCGTTTACATCATTTCCAAAATCTT 912
Db 940 CTCCACGCTTGTACAAACATTTTACAGGACATGGGTATTTGATCATCGTTCCAGAAATC 999
Qy 913 CTTGACAACTTTTCTGTTTCTCTTTTGTGAGTTACTATTTGATTCAGCTTTCACCCACAG 972
Db 1000 CTGGACAATATATATTCATTTCTCTCTTTTGAAGCCACGCTAGATCTCGATTTCCATCTCAG 1059
Qy 973 CTCCATGCTCTCTTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAGCA 1032
Db 1060 CTCCATGTTTGTGAGCAGGTTGTGGATTTGATTTGTTGTTGATGAAGCAAAACCT 1119
Qy 1033 GAAAGGGGTCCAAACAAAGCACATGCCACACCTGAAACAGTGGACCAATGTGTTCAACCCCT 1092
Db 1120 GAAAGAGCTCCCAACAAACACATGCCACCTCCAGCTCAATGGACTAAGCATTCATCTCT 1179
Qy 1093 GCATTTTCAATATATGATGCTACTGCTATGCTTAATTTTCAACCTTAAACAAAGCTGCT 1152
Db 1180 GCATTTTCTGATATGCTACTATTTGTTATGCTAACCTCTATGTTGTTTAAATTAAGCTCGA 1239
Qy 1153 GAGTCAAGGAGGATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTT 1212
Db 1240 GAGTCAAGGAGGATGACCTACTATCACTACGACCAACATTTCTGGAGAGGCTGGTGAAT 1299
```



```

QY 1213 GATCACTTGGCAGCGCATTTCTTCTCTGTGTCACACATATCATCGGAATTAATCTAAGG 1272
Db 1300 GACCCTTGGCTGTACGTTTCTAATGCCATAGATCGCATGGAATCAATCTGCGA 1359
QY 1273 AAGTCTCTGTGCTTCCAGTACTTGTACTATCTTGGTTCAGATTGGTCTGGCGATGTCGCCA 1332
Db 1360 AAGTCTCTGTGCTTCCAGTACTTGTACTATCTTGGTTCAGATTGGTCTGGCGATGTCGCCA 1419
QY 1333 TTGAGCAACAATCTCTTATTTCTTGACTATCATCTGCAACCCCTTTCCAAAGCTTCTTCCAA 1392
Db 1420 CTGAGCAACAATCTCTTATTTCTAGATTACCAACCGGAACCCGTTTCTGTGTTTCTTTA 1479
QY 1393 CGAGGTCGAATGCTCATATCTACGATGAGCCCTTTGCAAAATCACTGACAAAGAA 1452
Db 1480 AGAGGTCCTCAATGTTTCTCTGCTACTGATGACCCCTTCAGATTCACTTAATAAGAA 1539
QY 1453 CCAATGGTGGAGAAATACAGCATTTGCTGCTTGGTGGAAAGCTCAGTTCTTGTGATTTA 1512
Db 1540 CCTCTCTGGAGAGTATAGCATAGCTGCATCAGTTTGGAGCTGAGTGGTGTGACCTG 1599
QY 1513 TGGAAATTTGGAGGAATCTCTGTTTACCAATCTGGGTTTTCATATGCTCTCAAGGCGAC 1572
Db 1600 TGGAGATAGCTCGTAACCTCAGTGTACCACTCAGGTCTTCTCACACGCCCTGAAGTGGCAC 1659
QY 1573 TGGATTGTGAAGAACTACTTCAAAAGAGGACCTGCTGGAATGATATTCACAGAACCAAT 1632
Db 1660 TGGATTGTGAAGAACTACTTCAAAAGAGGACCTGATGGAACGACATTCACAAAACAAAC 1719
QY 1633 GTACCCACATCAGGGTTCATATTTAGAGATGATCTGGA 1672
Db 1720 GTGCCACATAGGGTGGAGTTCCTGTGACACGGTATGGA 1759

```

## RESULT 10

```

AX0811198
LOCUS AX0811198 2482 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 17 from Patent WO0109305.
ACCESSION AX0811198
VERSION AX0811198.1 GI:13170074
KEYWORDS Glycine max (soybean)
SOURCE Glycine max
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1
REFERENCE 1
AUTHORS Caspar, T., Falco, S. C., Sakai, H., Weng, Z. and Hu, X.
TITILE Purine metabolism genes in plants
JOURNAL Patent: WO 0109305-A 17 08-FEB-2001;
E.I. DU PONT DE NEUMOURS AND COMPANY (US) ; PIONEER HI-BRED
INTERNATIONAL, INC. (US)
FEATURES
source
1. .2482
/organism="Glycine max"
/mol_type="genomic DNA"
/db_xref="taxon:3847"
BASE COUNT 688 a 539 c 569 g 686 t
ORIGIN

```

Query Match 50.4%; Score 962.4; DB 6; Length 2482;

Best Local Similarity 78.7%; Pred. No. 3.3e-190;

Matches 1149; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

```

QY 13 GTTGCTCCATGGGAGGAGGTCATAAATGACCCCTGTACTCCAAACCTTAACCCCAAC 72
Db 1023 GTTGCTCCGTTGGATAAGAGTTATATCCGACCCGACACACCCAGGCTTAACCCAGAT 1082

```

```

QY 73 CCGTTCACTTATGTGCTGAAACCAAGTCAGAGATGTTTCCAAACTGTTGATGGGTT 132
Db 1083 CCATTTTATACATCTCTGAAGGAAATTCGTGATCATTTATTTGAAATGCAAGATGGGTT 1142

```

```

QY 133 ATCCATGTTTATGGGATAAAGATGTTACGGAGAGCATTTATCTCTGTGGCTGATGCTACA 192
Db 1143 ATTCTGTATATCAGATAGAGATCCAAAGAGAGCTTTTCTCTGTAGCCGATCAACT 1202
QY 193 ACCTTCTTCTGACTTGCATTTATTTCTCGAGTAAACGGCTGACGGAAACACAAGAACT 252
Db 1203 ACATTTTCTCACTGATCTTCACTTCTCGAGTCATAGCAGCAGGAGGAATATAAGAACT 1262
QY 253 GTCTGCCAATACGGTTAAATCTTCTTGAGCATAGTTTAAATCCATCTCATGTTTAAAT 312
Db 1263 TTATGCCATCATAGGCTCAATCTTCTAGAACAAAATCAATCTTCAITGATGCTAAAT 1322
QY 313 GCGGATAGGGAAATTTCTTGGCCAGAAAGACTGCCACATCGTGAATTTTACAATGCTCAGG 372
Db 1323 GCGGATAGAGAAATTTCTTGGCTCAGNAGAGTGTCCACATCGAGACTTCTATAATGTTAGA 1382
QY 373 AAGTTGACATCTATGTTTCATCTTCACTGACATGCAATCAAAAACATCTCTGTGAGGTTT 432
Db 1383 AAGTTGATCTCATGTCCACCTCAGCATGCAATCAGAAAACATCTTTTAAGGTTT 1442
QY 433 ATAAATCCAAACTAAGAAAAGAACCTGATGAGGTGCTCATTTTCAGAGATGGTACTTAT 492
Db 1443 ATAAAGTCAAAAGCTGAGAAAAGAGCCTGATGAGGTGTTAATATTTTCGAGATGGGACATAT 1502
QY 493 ATGACTTTAAAAGAGGTTTTCGAGAGCTTGAGACTTAACTGGGTATGATCTGAAATGTTGAT 552
Db 1503 CTAAGTTGGAGAGGTTTTCAGAGTTTAGATTGTTCTGGNATGACCTCAATCTTGAC 1562
QY 553 TTGCTAGATGTCATGACAGCAAAAGACATTTTCATCTGTTTGTGACAAATTCATCTAAAA 612
Db 1563 CTTTTCGAGCTTCAACGACAGAGAGTACTTTTCATCTGCTTGTAAAGTTCAATCTTAAA 1622
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTTCAACAGATATCTTAT 672
Db 1623 TACAATCTCTGCGSTCAAGTAGGCTCAGGGAGATATTTCTTAAAGACAGATATCTCAT 1682
QY 673 CAAGGCCGTTTCTTGTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAA 732
Db 1683 CAAGTCTGTTTCTTGTGAGTTAACTAAGCAAGTGTTCAGATCTTCTGCTGCGAGTAAA 1742
QY 733 TATCAGATGGCAAAATATAGGATTTCAATCTACGGAAGAAACACAGATGAATGGGACCAA 792
Db 1743 TATCAGATGGCTGAAATATAGAAATCAATATATGTTAGGAAAGCAAGTGAAGTGGACCAA 1802
QY 793 CTTGCAAGTTGGATAGTGAACAAATGCAACAGTGGAAATGTTGCTGCGCTGCTGAC 852
Db 1803 CTAGCCAGTTGGATAGTGAATATGATTTGTACAGGAGAAATGTCGTTGTTGTTGATTCAG 1862
QY 853 ATTCACGCTTATATAATGTGTACAAAGAAATGGGTATCGTTTACATCATTTCCAAAATCTT 912
Db 1863 CTTCCACGGTTGTACAATGTGTACAAAGAAATGGGAAATGTGACATCATTTCCAGAACATG 1922
QY 913 CTTGCAACATTTTCGTTCTCTCTTTTGGGTTTACTATTTGATCCAGCTTTCACACCCACAG 972
Db 1923 CTCGCAATATTTTCATTTCCACTTTTTCAGGTCCTGTCACCCAGATTCACATCTCTCAG 1982
QY 973 CTCATGCTTCTCTGACAGAGGTTGTAGGTTGGACCTGTTGATGATGAAGTAAACCA 1032
Db 1983 CTTGATGTTTCTCTGAAACAGGTTGTGGGTTTGGATTTGGTTGATGATGAAGAACCACT 2042
QY 1033 GAAAGCGCTCCAAACAAAGCACATGCCACACCTCGAACAGTGGACCAATGTGTTTCAACCT 1092
Db 2043 GAAAGCGGCCAAACAAACACATGCTACACCTGAGCAATGACTATATGTTTTCATCCG 2102
QY 1093 GCATTTTCAATATTTATGCGTACTTCTGCTATGCTAACTTATTCACCTTAAACAGTGGCT 1152
Db 2103 GCATTTTCACTACTGTCTATTACTGTTATGCAAAATCTTTACACCTTAAACAGCTTCA 2162
QY 1153 GAGTCAGAGGGAATGACCACTATCAATTTCCGTCACATGCTGGAGAGGCTGGAGATGTT 1212
Db 2163 GAATCAAGGGAATGACAAACAAATTTCCGTCACATCTCTGGAGAGGCTGGTATAT 2222
QY 1213 GATCACTTGGCAGCGACATTTCTTCTCTGTCTCAACATATCATCGGAATTAATCTAAGG 1272

```



```

RESULT 12
AX357098
LOCUS AX357098 914 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 17 from Patent WO0206319.
ACCESSION AX357098
VERSION AX357098.1 GI:186744279
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Schulz, A., Streber, W., Hanke, C., Schmidt, F. and Schubel, A.
Nucleic acid molecule which codes for a plant amp deaminase
Patent: WO 0206319-A 17 24-JAN-2002;
Aventis CropScience GmbH (DE)
Location/Qualifiers
1. .914
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
<2. .880
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD23277.1"
/db_xref="GI:18674280"
/translation="PRLYNIYKDMGIVTSFQNILNFIPLFEATVDPDSHPQLHVFL
KQVGFVLDVDES KPERRTKMPPTQAFWTFDPAFYSYVYVYCYANLYVNLKRESK
GWTITLRLPHSGEAGIDHLAATFLTCHSIAGHINLRKSPVLQYLYLYLAQIGLWMSPL
SNNSFLDVRHPFPVFFLRGLNLSLTDPLQIHLTKPLVEEYSIAASVWKLASCD
LCEIARNSVYQSGFSEHAKSHWIKGRDYKRGPDGNDIHKTNVPHIRVERFDITWKEEM
QQYLGKAVISDEVP"
BASE COUNT 250 a 215 c 196 g 253 t
ORIGIN
Query Match 27.8%; Score 531.6; DB 6; Length 914;
Best Local Similarity 76.6%; Pred. No. 1.7e-100;
Matches 651; Conservative 0; Mismatches 199; Indels 0; Gaps 0;
QY 856 CCACGCTTATATATGTGTACAAAGGAATGGGTATCGTTACATCATTCACAAATCTTCTT 915
DB 2 CCACGCTTGTACAAATTTACAAAGGACATGGGTATTGTGACATCGTTCCAGATAATCTCTG 61
QY 916 GACAAATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 975
DB 62 GACAATATATATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 121
QY 976 CATGTCTTCTCTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGATGATGATGATGATGAT 1035
DB 122 CATGTTTTTTTGAAGCAGGTTGTAGGTTGGATTTGATTTGATTTGATTTGATTTGATTTGAT 181
QY 1036 AGCGCTCCAAAGACACATGCCACACCTGAAACAGTGGACCAATGTGTTCAACCTCGCA 1095
DB 182 AGACGCTCCCAAAACACATGCCACCTCCAGCTCAATGAGTCAATGAGTCAATGAGTCAATG 241
QY 1096 TTTTCATATATCGGTACTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1155
DB 242 TTTTCGTTATATGCTACTATTGTTTATGCTTAACCTCTATGTTGTTAAATAGCTTTCGAGAG 301
QY 1156 TCAAGGGAATGACCACTATCAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTTGAT 1215
DB 302 TCAAGGGCATGACTACTATACGCTACGACCACTTCTGGAGAGGCTGGTGCATGATGAC 361
QY 1216 CACTTGGCAGCAGATTTCTTCTGTGTCAACAATATACATGGAATTAATCTTAAGGAAG 1275
DB 362 CACTTGGCTGCTAGCTTTCTTAACATGCCATAGCATCGCACATGGAATCAATCTGCGAAAG 421
QY 1276 TCTCCTGTGCTTCACTTGTACTATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1335
DB 422 TCTCCTGTGCTTCACTTGTACTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481

```

```

QY 1336 AGCAACAACTCCTTTATTTCTTGTGACTATCATCGCAACCCCTTTTCCAAACGTTCTTCCAACA 1395
DB 482 AGCAACAACTCCTTTGTTTCTAGATTACCAACGGAACCCGTTTCTCTGTCTTTTCTTAAGA 541
QY 1396 GGTCTGAATGCTCATTTATCTACGGATGACCCCTTTGGAAATTCACCTTCAACAAGAACCA 1455
DB 542 GGTCTCAATGTTTCTCTCTACTGATGACCCCTTTCAGATTCACTTGAAGAAACCT 601
QY 1456 TTGTTGGAAGAATACAGCAATTCGCTTGGTGGAGAGCTCAGTTCTTGTGATTATATGC 1515
DB 602 CTCGTGGGAAGATATAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
QY 1516 GAAATTGCGAGGAACCTCTGTTTACCAATCTCGGTTTTCACATGCTCTCAAGGCGCAGCTGG 1575
DB 662 GAGATAGCTCGTAACTCAGTGTACAGTCAAGTCTTCTCACACGCCCTGAAGTCGCACTGG 721
QY 1576 ATTGTTAGAACTACTTCAAAAGAGGACCTCTGGAATGATATTCACAGAACCAATGTA 1635
DB 722 ATTGAAAAGATTACTACAAAAGAGGACCTGATGAAAACGACATTCACAAAACCAACCTGG 781
QY 1636 CGGCACATCAGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAAATGAAACTAGTGTAC 1695
DB 782 CCACACATAAGGTTGGAGTTCGTGACACAGATCTGGAAGAGGAGATGCAACAGGTTTAT 841
QY 1696 TCTGACAATG 1705
DB 842 CTGGGCAAG 851

SPBC106 42544 bp DNA linear PLN 08-DEC-1999
S. pombe chromosome II cosmid c106.
AL110295
AL110295.2 GI:6562182
60S ribosomal protein 125-a; adal; Adenosine/AMPdeaminase; amp
deaminase; camp-dependent protein kinase; cct4; chaperonin;
chaperonin subunit cct4; cut4; exocyst complex; homoserine
o-acetyltransferase; idl1; isopentenyl-diphosphate delta-isomerase;
mph1; pkai; PROS28 family; proteasome component; protein kinase
homolog; RNA binding protein; rpi25a; tpx.
Schizosaccharomyces pombe (fission yeast)
Schizosaccharomyces pombe
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
1 (bases 1 to 42544)
Lyne, M., Rajandream, M.A., Barrell, B.G., Davis, P. and Churcher, C.M.
Submitted (26-AUG-1999) European Schizosaccharomycetes genome
sequencing project, Sanger Centre, The Wellcome Trust Genome
Campus, Hinxton, Cambridge CB10 1SA, E-mail: barr@anger.ac.uk
On Dec 12, 1999 this sequence version replaced gi:5817266.
Notes:
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/S_pombe/)
During 1995 to 1996 about 66% of S. pombe chromosome 1 was
sequenced by the Sanger Centre. The sequencing of the S. pombe
genome is now being continued with funding from The European
Commission. Fourteen European sequencing laboratories, including
the Sanger Centre, are participating in the project.
Protein coding regions (CDS) have been predicted with the help of
computer analysis using the Genefinder program in PomBase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPBC25H2.01c. SP (S.
pombe), B (chromosome 2), c25H2 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database

```

are also included but some of these may be fortuitous.  
The length in codons is given for each CDS.  
IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.  
Cosmid c106 is overlapped at the 3' end by cosmid c582, EMBL entry SPBC582, accession number AL096788.

## FEATURES

```

source
    Location/Qualifiers
        1..4254
        /organism="Schizosaccharomyces pombe"
        /mol_type="genomic DNA"
        /strain="972h-"
        /db_xref="taxon:4896"
        /chromosome="II"
        /map="11L"
        /clone="cosmid c106"
gene
    1..282
    /gene="SPBC106.01"
    /note="synonym: mph1"
    1..282
    /partial
    /label=mph1
    /product="protein kinase homolog"
    /protein_id="CAB53717.1"
    /db_xref="GI:5817267"
    /translation="RDQRKLTIPPELLVHPPLNPVSYLTPLAKKPLPVSGHTNNAHP
    LRLSTEIASQSLMIERSVLSKHLINKELIDSMAYDCVSNLRKPE"
    Complement(819..1193)
    /gene="SPBC106.02c"
    complement(819..1193)
    /note="SPBC106.02c"
    /gene="SPBC106.02c"
    /note="SPBC106.02c, len:124, SIMILARITY: Saccharomyces
    cerevisiae, YK16_YEAST, hypothetical 13.9 kd protein in
    cyt2-mdh1 intergenic region., (127 aa), fasta scores: opt:
    317, E(): 8.1e-16, (45.1% identity in 113 aa)"
    /codon_start=1
    /label=SPBC106.02c
    /product="hypothetical protein"
    /protein_id="CAB53718.1"
    /db_xref="GI:5817268"
    /translation="MTSHTGNNNNIVELDMSELIRPIPPVLDMMKNVNSMMETMTGKT
    PASCGLTSEDLGELPPVDVLTFFKSGKPYFAGGCHRLRAHDEAGRKKVRCKLV
    NCSPNTLRLLGASANKFLSD"
    2984..4057
    /gene="SPBC106.03"
    /note="SPBC106.03"
    /note="SPBC106.03, len:357, SIMILARITY: LOW TO
    Saccharomyces douglasii, YH7_SACDO, hypothetical 44.4 kd
    protein in sp013-arg4 intergenic region., (385 aa), fasta
    scores: opt: 122, E(): 0.47, (24.0% identity in 287 aa)"
    /codon_start=1
    /label=SPBC106.03
    /product="hypothetical protein"
    /protein_id="CAB53719.1"
    /db_xref="GI:5817269"
    /translation="MNPERRAADLVSNVYDKRIIDVAHEKTSNPQPVTVHTNSAK
    GRLEFLYONVRVGLSAAFLPLAGLSAYYKRYDYVRKRAPRAPHNQKTVVLSAWN
    PLASYADHLDRGFIYVVLVDSAEATYVLRQRPYQTSILVTHNEAVERFLSFSN
    QSGPEYALRLGLIIVPTGSLVPTPIENIGKDAWISAFQOLSESFNSVSMIPLKLS
    QKARIIGLSHGLSAYEPNPVAPVLSLSSTIETPLRLKRETLQVLCIKLGNLSFLN
    YDHSSEKSNYPSPSLCTERKVLNKFPSILGCVPSFRRHVGCRFFFMVTVSRFLPTC
    VIDGIFSLGKHFHFCSLIKRS"
    Join(4585...4622,4740..4917,4981..7305)
    /gene="SPBC106.04"
    /note="synonym: adal"
    Join(4585..4622,4740..4917,4981..7305)
    /gene="SPBC106.04"
    /note="SPBC106.04, len:846"

```

```

/codon_start=1
/label=adal
/product="amp deaminase"
/protein_id="CAB53720.1"
/db_xref="GI:5817270"
/translation="MSVRPLSLCGLLPAEMNMEQDDQVPAVAETVPLKRYVTNPGA
NRDEVAARASSQTPFYDYAYERSLHQDAKFLAMNGTQNGRGLSKSPRRFSVSA
STVNSDDVNSKAGSGKLLNDLTQSKISSIHMPHQGDNAVSVSGPEDTDPGN
METDPLFSDLAELIYLSIHKMDMRHYIRVSLQGLDNPIDDDSIYIPDCKEED
DTGLFNADCKIPGIENEMEHMDHQIFQVYENDSAYIAGTSPFHPTIRDYVIDLE
FLLSASDGPSSKSPFRRLQVLEGRWNMYMLNEYQELADTKVPHRDFYNNRVKVDTH
VHLSALANOKHELLRFIKAKRKCPNKXVIRWDGKFLTQEVFDSLKLTSDLSIDTLDY
MHAUTDFRDKFNKNTNPIGESRLRTIFLKTNDINDINGRLABLTKREVFTDLRTQKY
QMAEYRISIYGRNRREWDKLAAMIDNELFSPNVRLIQVRLYDVYKKGIVETFE
VNRVFEPLFEVTRKPRTHKLVFLQGVIGFSDVDESKPERTFRRKFPYKHWIDIN
LNPYSYWLVMYAMNTSLNSKIRGFNTFVLPHGCEAGDTDLHLSAFLLSHGINH
GILLRKVFLQYLWLDQIPIAMSPISNNALFLAYDKNPFLTYPKRGVLNLSLTDPL
QFATREPLIEEYAAQIKLSAIVDCMLARNSVLOSGEROLKEKWLGVDFQDIDR
TNVFIIRLAYRALTLTQELIALNVKHVQSPKSPSHDLIELIHKDYDMTGTSDPLSASP
RTNDATISSRLSLHDGHDGFAFFGLSVISERRRRKDSMASSQDLKD"
4623..4628
/gene="SPBC106.04"
/note="gtatat, splice donor sequence"
4729..4739
/gene="SPBC106.04"
/note="ctaatatttag, splice branch and acceptor"
4918..4980
/gene="SPBC106.04"
/note="confirmed intron"
4918..4923
/gene="SPBC106.04"
/note="gttaagt, splice donor sequence"
4966..4980
/gene="SPBC106.04"
/note="ctaacataacctag, splice branch and acceptor"
6178..6963
/gene="SPBC106.04"
/note="Match to PF00962 A_deaminase, Adenosine/AMP
deaminase Score 419.88"
complement(join(7884..7969,8060..8189))
/gene="SPBC106.05c"
complement(join(7884..7969,8060..8189))
/gene="SPBC106.05c"
/codon_start=1
/label=SPBC106.05c
/product="hypothetical protein"
/protein_id="CAB53721.1"
/db_xref="GI:5817271"
/translation="MYEHRSHIQCKQKQKQKIDENYHROESLIESAKIAYLNTKNTPPKE
DSMLPNLKDSDFDMDDEFVKELEKNV"
complement(7970..7985)
/gene="SPBC106.05c"
/note="ctaatatacaatttag, splice branch and acceptor"
complement(8054..8059)
/gene="SPBC106.05c"
/note="gttagt, splice donor sequence"
8678..10261
/gene="SPBC106.06"
/note="synonym: cct4"
8678..10261
/gene="SPBC106.06"
/note="SPBC106.06, len:527"
/codon_start=1
/label=cct4
/product="chaperonin subunit cct4"
/protein_id="CAB53722.1"
/db_xref="GI:5817272"
/translation="MSKAATVPVAFODREKPEQVRLSNIAARSAVDARTSLGPKGM
DKMLOTGKGEVILTNDGATILKHLVHPAAKMLVDLSAQDYVAGDGTTSVILAGS
MLCAEKLLKKGHIPTVIAESFORAAGFTVDCMKENALAIELSDRELLRAATSLNS
KIVSQSNLAPIAVDAVLKVIDPVRATNVNLDKDIRIVKGLGGIIDDTLIPGLALTQ
TAVKAGGPTRIEKANIALIQFQLSPPKPMENQVNVVNDYROMDKILKEERQVLLNMC

```

KKIKAGANVILIQSIILRDVNDLALHPLAKLIMVINKIDRDEVEFICKSTGCKPI  
ADIEFAEDKHADLVBETSSGEKIVKFSGVKNAGKTVSILCRGANLLTLEAEARS  
LHDALCVRLVKORALJAGGSPLEBAORLLEHAROLEGFEALICIRAFSEALBIP  
VTLAENAGNALQVOTELSRHANGKTAGINVRKGIIVTNILLENVLQBLUNISAIQ  
LAAETTKMIMKIDIDITLAR"

misc\_feature

8768..10246

/gene="SPBC106.06"  
/note="Match to PF00118 cpn60\_TCP1, TCP-1/cpn60 chaperonin  
family Score 634.60"  
complement(10863..11366)  
/gene="SPBC106.07c"  
complement(10863..11366)

CDS

Query Match 26.1%; Score 498.6; DB 8; Length 42544;  
Best Local Similarity 60.1%; Pred. No. 1.1e-93;  
Matches 828; Conservative 0; Mismatches 549; Indels 0; Gaps 0;

QY 205 GACTTCGATATATATCTCGAGTAACGGCTGCAGGGAACACAGAACTGCTGCCATAT 264  
DB 5584 GATTTAGAATTCCTCTTCGGCTTCATCGGATGGTCTAGTAATCAATTTTCCTTCGCT 5643

QY 265 CGGTTAAATCTCTTGAGCATAGTTTAAATCCATCTCATGTTAAATCGGATAGGAA 324  
DB 5644 AGATTGCAGTATCTAGAAGCAGCTGGAATATATATCTGCTAAACGAGTATCAAGAG 5703

QY 325 TTCTTGGCCAGAGACTGCCACATCGTGATTTTACAATGTCAGGAAGGTTGACACT 384  
DB 5704 CTGGCAGATACCCAAAAGTACCCCATCGTGATTTTCAATGCTAGTAAGGTAGTACT 5763

QY 385 CATGTTTCATTCAGCATGCATGAATCAAAAACATCTGTGTAGGTTTCATAAATCCAAA 444  
DB 5764 CAGTCCATCATCTCTCTGGAATCAAAAACATCTCTTGGCATTTATCAAGCTAAA 5823

QY 445 CTAAGAAAGAACCTGATGAGGTGGTCATTTTCAGAGATGGTACTTATATGACTTTAAA 504  
DB 5824 CTTTCGTAATGCCTTAATGAAAAGTTATCTGCGGTGATGGCAATTTCTGACTTTTACAA 5883

QY 505 GAGGTTTTGAGAGCTTGACTTAACCTGGGTATGATCTGAATGTTGATTTGCTAGATGTC 564  
DB 5884 GAAGTATTTGATTCGTTGAAGTTGACTTCGTATGACCTTTAGTATAGACAGCTTGGATATG 5943

QY 565 CATGCAGACAAAAGCATTTCATCGTTTTTGACAAAATTCAAATCAAAATCAATCCATGT 624  
DB 5944 CATGCTCATACTGACACCTTTCCACGATTTTGACAAAATTTAACTTGAATATATAATCAAT 6003

QY 625 GGCAGAAAGTAGGCTCAGAGAAATTTTCTCAACACAGATATCTTATTCAGGCCGTTTT 584  
DB 6004 GGTGAATCTCGTTTGGCTACTATTTTCTGAAAACCGGATAATGATATTAACGCTCGATAT 6063

QY 685 CTTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTCTGCTAGCAAAATATCAGATGGCA 744  
DB 6064 CTTGCTGAACTTACCAAGAGGTGTTTACCGACCTTTAGAACTCAAAAGTACCAATGGCT 6123

QY 745 GAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTGCAGAGTTG 804  
DB 6124 GAATATCGTATTTCTATCTATGACGGAATCGTGAAGATGGGTAAGCTTGACGCTGG 6183

QY 805 ATAGTGACAAATGAATGACAGTGGAAATGTTGCTGGCTGGTTGAGATTCACACGCTTA 864  
DB 6184 ATTATTGATAACGAACCTTTTTTCTCCCAACGTTTCTGGTGTATTAAGTACCTCGTTTG 6243

QY 865 TATAATGTGTACAAAGGAATGGGTATGTTTACATCATCTCCAAATCTTCTGTGACACATT 924  
DB 6244 TATGATGTGTATGAAGATCGGTTATTTGAGACTTTTGAAGAGGTGCTGCAGAAATGTC 6303

QY 925 TTGCTTCTCTTTTGGAGTTACTTATGTATCCAGCTTTCACACCAAGCTTCCATGCTTC 984  
DB 6304 TTTGAACCACTGTTTGAAGTTTACAAAGGATCTCTGACACATCTTAACTCCACGCTTT 6363

QY 985 CTGAAGAGGTTGTAGGTTGGACCTGGTTGTATGATGAAGTAAACAGCAAGGCGTCCA 1044  
DB 6364 TTGCAACGTGTTATCGGTTTGTACTCTGTTGACGACGAAAGTAAACCTGAACCCGTTACC 6423

QY 1045 ACAAGCACATGCCACACACCTGAAACAGTGGACCAATGTGTTCAACCTCGCATTTTCATAT 1104  
DB 6424 TTTCCGAAATTCCTTATCCGAAACATTTGGGATATTAATTTAAATTCCTCCATAGTTAT 6483  
QY 1105 TATGGTACTACTGCTATGCTAACTTATTCACCCCTAAACAGCTCGCTGAGTCAAGGGA 1164  
DB 6484 TGGTTATATTACATGTTATGCCAACATGACTAGCTTGAATAGCTGGAGAAAAATACGTGGT 6543  
QY 1165 ATGACCACATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTTGTACACTTGGCA 1224  
DB 6544 TTTAAACACTTTGCTTCCTCGTCCATGTTGGTGAAGCTGGAGATACGGATCAITTTGGCC 6603  
QY 1225 GCGACATTTCTCTGTCAACAACATATCATGGAATTAATCTAAGGAAGTCTCCTGTG 1284  
DB 6604 AGTCTTTCTTATTAAAGTCACGGTATTAAACATGATTTTATTACGAAAAAGTTCTCTTT 6663  
QY 1285 CTTGAGTACTTGTACTATCTTTGGTCAGATTGGTCTGGGATGTCCTCCATTCAGCAACAC 1344  
DB 6664 CTTGAGTACTTGTACTATCTTTGGTCAGATTGGTCTGGGATGTCCTCCATTCAGCAACAC 6723  
QY 1345 TCCTTATTTCTGACTATCATCGCAACCTTTTCCACGTTTCTTCCACAGAGGTCTGAAT 1404  
DB 6724 GCACTGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6783  
QY 1405 GTCTCATTTATCTACGATGACCTTTGCAAAATTCACCTGACAAAAGAACCTTGGTGAA 1464  
DB 6784 GTTAGCTTTCCACTGATGATCTTTTACAAATTTGCAATTTACCCGAGAGCTCTGATTGAG 6843  
QY 1465 GAATACAGATGCTGCTTCTGCTGGAGCTCAGTTCTTGTGATTTATGCAAAATTCGG 1524  
DB 6844 GAATATGCTGTAGCGGCAAAATTTATAAGTCTCTGCTGCGATATGTCGAATTTGGCT 6903  
QY 1525 AGGAACCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCACTGGATTCGT 1581  
DB 6904 AGAAACAGTGTTHACATCCGGTTTCGAAAGTCAAGCTTAAGAACGTTGGTTGGT 6960

RESULT 14

BC049119 3667 bp mRNA linear ROD 22-APR-2003  
Mus musculus adenosine monophosphate deaminase 2 (isoform L), mRNA  
(cDNA clone MGC:61170 IMAGE:6812571), complete cds.

BC049119

ACCESSION

VERSION

BC049119.1 GI:29145072

KEYWORDS

MGC

Mus musculus

(house mouse)

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 3667)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,  
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,  
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,  
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,  
Diachenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,  
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,  
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,  
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,  
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,  
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,  
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,  
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,  
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,  
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,  
Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,  
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,  
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalls, D.E.,  
Schnerch, A., Schein, J.E., Jones, S.I. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
22388257  
PUBMED 12477932

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 3667)  
Strausberg,R.  
Direct Submission  
Submitted (19-MAR-2003) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

REMARK  
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Dr. Jim Lin, University of Iowa  
cDNA Library Preparation: M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.  
Thomas L. Casavant.  
Web site: <http://genome.uiowa.edu>  
Contact: [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu); [tom-casavant@uiowa.edu](mailto:tom-casavant@uiowa.edu)  
Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,  
Fisher,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,  
Sheetz,F., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,  
Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: Plate: Row: Column: 0.

FEATURES  
source

1..3667  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6"  
/db\_xref="taxon:10090"  
/clone="MGC:61170 IMAGE:6812571"  
/tissue\_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"  
/clone\_lib="NIH BMAP\_FW0"  
/lab\_host="DH10B"  
/note="Vector: pYX-ASC"

## gene

1..3667  
/gene="Ampd2"  
/note="synonyms: Ampd-2, 1200014F01rik"  
/db\_xref="LocusID:109674"  
/db\_xref="MGI:88016"

## CDS

401..2797  
/codon\_start=1  
/product="Ampd2 protein"  
/protein\_id="AAH49119.1"  
/db\_xref="GI:29145073"  
/db\_xref="LocusID:109674"  
/translation="MASEARSLGASPLQASRLPGNAPCLKHPFLDLRTSMGKCKE  
IAELFSLARSLERAPYEFPEESIEQLEERRQRLERISQDVKLEPDLILRAKQ  
DLKTDSDSLQLYKEQEGQDGLWEDVVLREFORVILSGEKGVPFTDLDA  
AKSVVRLFIREKYMALSLQSPTRRYLQQLAEKPLETRTYEQSPDTPVSADAPVH  
PPALQHPYHECPGAMPDGLGLRMVGVVHVRDPDEHCPVELPVPDLQEPV  
ADVNLMLIINGPIKSCYRLOVLSKFQMHVLLNEMKELAAQKVPKPHRDFYNIRK  
VDTHIHASCMNQKHLRPIKRAMKHLEIEIVHVEQREQTILREVFESEMLITADLSV  
DILDVHARLTHREDKFNKPNPAGESVLRIFIKTDNKSQKFAHIKEVWADLE  
ESKYQNALRSLISGRDEWDKLARVNVHVKHSPNVRLVQVPRFDVYRTKGOLA  
NFQEMLENIPLFEPAPHPHELHFLFLEHVDGDFSDDESKPENHVFNLSPLE  
AWVEDNPYAYLYVTANMAMNLHRLRQGFHTFVLRPHCGEAGPIHILVSAFMLA  
ENISHLGLLRKAPVLYLYLAQIGIAMSPLSNLSFLSYHNPPLPEYLSGLMVLSL  
TDDPLQFHTKEPLNEEVSIAQVWKLSCDMCLARNLSVMSGFSHKVKHSLGPNY  
TKEGPEGNDIRKTPVDIRGVYRTGCTTCCATACCTGACCTACAGNAITTTAGCTG 1232

BASE COUNT  
ORIGIN

759 a 1051 c 1030 g 827 t  
Query Match 25.5%; Score 487.8; DB 10; Length 3667;  
Best Local Similarity 58.2%; Pred. No. 2.1e-91;  
Matches 877; Conservative 0; Mismatches 627; Indels 3; Gaps 1;

Qy 146 CGGATAAGATTGTACGGAGAGCAATTTATCCCTGGCTGATGCTACAACTTTCTTCACTG 205  
Db 1173 CTGATGAGCACTGTCCGGAGGTGGAGCTTCCATACCTGACCTACAGNAITTTAGCTG 1232

Qy 206 ACTTGCAATTATTTCTCGAGTAACGGCTGAGGGAAACAAGAACTGCTCTGCCATAATC 265  
Db 1233 ACGTCAATGTGCTGATGCCCTGATCATCAATGTGTCCCAATAGTCATTTCTGCTACCGCC 1292  
Qy 266 GGTAAATCTTCTTGAGCATAAGTTTAAATCCCATCTGATCTTAAATCCGATAGGGAAT 325  
Db 1293 GGCTGAGTACTCTGAGCTCCAAATTCAGATGCACGTTTTTCTCAATGAGATGAAGGAGC 1352  
Qy 326 TTCTTGCCAGAGAAGTGCCTCCACATCGTGATTTTTTACAATGTGAGGAAGGTGTGACATC 385  
Db 1353 TCGTGTCTCAGAGAAGTGCACACCGGACTTCTACAATATCCGTAAGGTGGACACAC 1412  
Qy 386 ATGTTTCATCTCAGCATGCAATCAAAACATCTGTTGAGGTTTCAAAAATCCAAAC 445  
Db 1413 ACATCCAGCGCTCGTCTGTCATGAACCAAGAAACATCTACTCGCTTCTATCAAGGGGCA 1472  
Qy 446 TAAGAAAAGAACCTGATGAGGTGTCATTTTTCAGAGATGGTACTTATATGACTTTAAAG 505  
Db 1473 TGAAGCGGCACCTGGAGGAGATTGTGCAATGGAAACAGGCGCGGAGCAGACGCTGAGAG 1532  
Qy 506 AGGTTTTTGAGAGCTTGGACTTAACTGGGTATGATCTGAATGTTGATTTGCTAGATGTCC 565  
Db 1533 AAGTCTTCGAGAGCATGAACCTCACTGCTACGACTTAAGTGTGACACGCTGGATGTGC 1592  
Qy 566 ATGACAGCAAAAGACATTTTCATCGTTTGAATAATCAATCTAAATPACAAATCCATGTG 625  
Db 1593 ATGCGGACAGGATACCTTTTCATCGAITGACAAATCAATGCCAAATACAAACCTATTG 1652  
Qy 626 GCCAAGTAGCTCAGAGAAATTTTCTCAACCAAGATAATCTTATTAACAGGCGTTTTTC 685  
Db 1653 GGGAGTCTGTTCTCCGAGAGATCTTCATTAACACCGACAAACAAGATTTCTGGGAAGTACT 1712  
Qy 686 TTGCTGAGTTGACAAAAGCAAGTTTTTCTCTGACCTTTCTGTAGCAAAATATCAGATGGCAG 745  
Db 1713 TTGCTCAGATCATCAAGGAGGTGATGCGAGACTTGGAGGAGCAATACAGATGAGAG 1772  
Qy 746 AATATAGATTTCATCTACGAAAGGAAACAAGATGAATGGGACAACTTCCAAAGTTGGA 805  
Db 1773 AGCTCCGCTGTCCATCTACGCGGCTTCGAGGATGAGTGGGACAAAGTGGACGCTGGG 1832  
Qy 806 TAGTGAAACAATGAAATTCACAGTGGAAATGTGCTCGCTGGTTTCAGATTCACAGCTTAT 865  
Db 1833 CAGTGAACCAACAAGTGCACCTCTCCCAATGTCCTGCTGGTGGTGGAGTGGCGGCTCT 1892  
Qy 866 ATAATGTGTAACAGGAAATGGGTATCGTTACATCATTTCCAAAATCTCTTTCGACAACTTT 925  
Db 1893 TCGATGTGACCGCACCAAGGCCAGCTGGCCAATCTCCAAGAGATGCTGAGAAACATCT 1952  
Qy 926 TCGTTCTCTTTTGGAGTTACTATTGATCCAGTTTCAACCCACAGCTCCATGCTTTCC 985  
Db 1953 TTCTGCCCCCTGTTTGGAGCTACTGTGCACTGCACTGCGAGCACCCGAGGCTGCACCTCTTC 2012  
Qy 986 TGAAGCAGTTGTAGGTTGGACCTGTTGATGATGAAGTAAACCAAGAAAGGCGTCCAA 1045  
Db 2013 TGGAGCAGTGGATGGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2072  
Qy 1046 CAAA---GCATGCCACACCTCAAAGTGGACCAATGTGTTCAACCTCCATTTTTCAT 1102  
Db 2073 TCAACCTGGAGAGTCCCTCCAGAAAGTGGGTGGAGGAGGACAACTCCCTCCCTATGCT 2132  
Qy 1103 ATTAGCGTACTACTGCTATGCTAACTTATTCACCTTAAACAAAGCTCGTGGTGAAGCAAAG 1162  
Db 2133 ACTACCTGTACTACACCTTGCTAACATGGCTATGTTGAACCATCTGCGCAGGAGAGAG 2192  
Qy 1163 GAATGACCATATCAAAATTCGTCCTCCATGCTGAGAGGCTGGAGATGTTGATCATTGG 1222  
Db 2193 GTTTCCACAGCTTCGTCGTGAGCGCCCACTGTGGGAGGCGCGGCCCATCCACACCTGG 2252  
Qy 1223 CAGCGCAATTTCTCTCTGTCAACAATATCACATGGAATTAATCTAAGGAAGTCTCTCTG 1282  
Db 2253 TATCAGCTTCATGCTGGCCGAGAAACATCTCCACAGGGGTGCTCTCTGGCAGAGGCCCG 2312  
Qy 1283 TGCTTCAGTACTGTACTATCTTGTGTCAGATTGCTGTCGAGATGCTGCCCAATTGAGCAACA 1342

```

Db      2313  TCCTGCAGTACCTGTATTACCTGGCTCAGATGGCATGCCATGTCCCGCTCAGCAACA 2372
Qy      1343  ACTCCATTATCTTGATCATATCGCAACCGCTTTTCCAAAGGTTCTTCCAAAGAGGTCTGA 1402
Db      2373  ACAGCCGTGTTCTTCAGTACATCCACCGAAGCCCTCTCCCTGAGTACTTGTCCCGTGGCCTCA 2432
Qy      1403  ATGTCTCATTTATCTACGGATGACCCCTTGCAAAATTCACCTGCACAAAGAACCAATTCGGTGG 1462
Db      2433  TGTGCTCGCTGTCCAGATGATCCCTTGAGTTCACCTTCACCAAGGAGCCCTGATGG 2492
Qy      1463  AAGAATACAGCATTTGCTGCTGCTGGAAGCTCAGTTCTTGTGATTTATGCGAAATG 1522
Db      2493  AGGAGTACAGCATCGCCACCGAGGTGGAAGCTCAGCTCTGCGATATGTCGAGCTGG 2552
Qy      1523  CGAGGAACCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGGGCGCATGGATGGTA 1582
Db      2553  CCGGTAAACGCGTCTCATGAGTGGCTTCTCTCAAGGTGAAAGCCATGCTGGCGAC 2612
Qy      1583  AGAATCTACTTCAAGAGGACCTGCTGGAATGATATTCACAGAACCAATGTACCGCACA 1642
Db      2613  CCAACTATACCAAGGAGGCCCTGAGGGCAATGATATCCGCGCTACCAAGCTGCCAGACA 2672
Qy      1643  TCAGGGT 1649
Db      2673  TCCGAGT 2679

```

## RESULT 15

```

SPADA1TCP
LOCUS      S. pombe adal and partial tcp1 genes.
DEFINITION X91498
ACCESSION  X91498
VERSION    1
KEYWORDS   adal gene; AMP deaminase; TCP1 gene.
SOURCE     Schizosaccharomyces pombe (fission yeast)
ORGANISM   Schizosaccharomycetes;
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomycetes.
REFERENCE  1
AUTHORS    Rochet, M., Levesque, H. and Gaillardin, C.
TITLE      Putative AMP deaminase in S. pombe
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 4761)
AUTHORS    Nicaud, J.M.
TITLE      Direct Submission
JOURNAL    Submitted (14-SEP-1995) J.M. Nicaud, Institut National de la
            Recherche Agron., Lab. Genetique Moleculaire et Cellulaire,
            Thiverval-Grignon, 78850, FRANCE

```

## FEATURES

```

source     Location/Qualifiers
            1..4761
             /organism="Schizosaccharomyces pombe"
             /mol_type="genomic DNA"
             /strain="972 h-"
             /db_xref="taxon:4896"
             /chromosome="II"
             /map="left arm"
             /clone="pMRSC1"
             /gene="adal"
             /db_xref="GI:995562"
             /db_xref="SWISS-PROT:P50998"
             /translation="MNGTQNGRLPSKPRRVSASVTNRSDVNHKAGPGSKL
             LNDLQSKISIRVHQGVQDNVSVSGFETDPGNMETTFDFDELAELIYLSHK
             CMDMRHKYIRVSLQGLNDNPDDSWIIYDPCKEGEDDGLFNFAKCKPIGNIENHE
             HMDHQGIFQVYENDSAYIAGTPSPHPTIRDYIDLEFLLSASDGPSKSFSPRLQY
             LEGRWNMYMLLNEYQELADTKVPHRDFYNVRKVDTHVHHSALANQKHLRFLKAKLR

```

## gene

## CDS

```

KCPNKEVIWRDGEKLTQEVDFSLKLTYSYDLSIDTLDHMAHTDTFHRFDKFNLYKNI
GESLRTIFLTDNDINGRYLAELTKVFTDLRTQKYQMAEYRISYGRNEEEDDKLA
AWIDNELFSPFTVGLPKYLKVMCMCIRSPVLLRLKRSEMSLNHSCFTKDPDRTHPK
LHVLQRIVGSDVDDSKSPERRTFKFPFKHNDINLPYSYKWLXYMYAMNLSLNS
WRKIRGNFTFVLRPHCGEAGDTDLHSAFLSLSHGINHGILLRKRVFFLOYLYWLOIPI
AMSPLSNNALFLAYDKNPFLLTYFKRGLNLSLSTDDPLQFAFTREPLIEEYAAQAIYK
LSAVDMCELARNSVLQSGFERQLKRWLVGFQDIDRTNVP IIRLAYRALTLTQSIAL
VKNHVQSPKPSNHDLEELIHKYDAMTGTSDPLSAPRNTDATTSSRLSLHDGHDGA
RFPGLSVISERRRRKDSMASSSQDLKD"
4456..4761
/gene="tcp1"
4456..4761
/gene="tcp1"
/codon_start=1
/protein_id="CAA62798.1"
/db_xref="GI:995563"
/db_xref="SWISS-PROT:P50998"
/translation="MSKAATVPVAFQDREKPEQVRLSNIMAARSVADAIKRTSLGPKGM
DKMIQTGKGEVILTNDGATILKHLVLPAAKMLVDLSAAQDVEAGDGTTSVVLGAS
"

BASE COUNT      1453 a      884 c      906 g      1518 t

Query Match      25.0%; Score 477.2; DB 8; Length 4761;
Best Local Similarity 60.1%; Pred. No. 3.3e-89;
Matches 828; Conservative 0; Mismatches 548; Indels 2; Gaps 2;

Qy      205  GACTTGCAATATATTTCTCGAGTAACGGCTGCGAGGAACACACAGAACACTGTCGCCATAAT 264
Db      1362  GATTTAGAATTCCTTTCTCGGCTTCATCGGATGGTCCTAGTAAATCATATTTCTCCGT 1421
Qy      265  CGGTTAAATCTCTTGAGCATAAGTTTAAATTCACATCTGATGTTAAATGCGGATAGGAA 324
Db      1422  AGATTGCAATCTTAGAAGGCGCTGGAAATATGTAATGCTGCTAAACGAGTATCAAG 1481
Qy      325  TTTCTTGCCGAGAAGACTGCCCAACATCGTGAATTTTACAATGTGAGAAAGTTGCACACT 384
Db      1482  CTGGCAGATACCAAAAAAGTACCCCATCGTGATTTTACAATGTACGTAAGGTAGATACT 1541
Qy      385  CATGTTTCATTCAGCATGCATGAATCAAAAAACATCTGTTGAGGTTTCAAAAAATCCAAA 444
Db      1542  CACGTCCATCATTTCTGCTTTGCGAATCAAAAAACATCTCTTGCAGTTTATCAAGCTAAA 1601
Qy      445  CTAAGAAAAGAACCTGATGAGTGCTCATTTTCAGAGATGGTACTTATATGACTTTAAAA 504
Db      1602  CTTCTAAATGCCCTTAATGAAAAAGTTATCTGGCGTATGGCAATTCCTGACTTTACAA 1661
Qy      505  GAGGTTTTGAGAGCTTGGACTTAACCTGGGTATGATCTGAAATGTTGATTTGCTAGATGC 564
Db      1662  GAAGTATTTGATTCTGTTGAAGTTGACTTCGTATGACCTTAGTATAGACACGTTGGATAG 1721
Qy      565  CATGAGACAAAAGCACATTTTCATCGTTTGAACAATTCATCTAAAATACATCCATGT 624
Db      1722  CATGCTCATCTACTGACACCTTTTCCCGAATTTGACAAATTTAACTTGAAATATATAAT 1781
Qy      625  GCCAAAAGTAGGCTCAGAGAAATTTTCTCAAAACAAGATAATCTTATTCAGGCGCTTT 684
Db      1782  GGTGAATCTCGTTTGGTACTATTTTCTGAAAACCGATATGATATTAACGGTCGATAT 1841
Qy      685  CTTGCTGAGTTGACAAAAGCAAGTTTTTCTCTGACCTTTCTGCTAGCAAAATATCAGATG 744
Db      1842  CTTGCTGAACTTACCAAGAGGTTTCCCGACCTTTAGAACTCAAAAGTACCAAAATGGCT 1901
Qy      745  GAATATAGGATTTCAATCTACGGAAGGAAAACAGATGAATGGGACCACTTCCAGATTGG 804
Db      1902  GAATATCGTATTTCTATCTATGACGGAATCGTGAAGAACGTAAGCTTCAGCGCTGG 1961
Qy      805  ATAGTGAACAATGAATGACAGTGGAAATGTTGCTGGCTGTTGCTGAGTTCAGATTCACG 864
Db      1962  ATTATTGATACGAACCTTTTTTCT -CCACAGTTGCTTGGCTTATTCAGTACTCTGTTG 2020
Qy      865  TATAATGTGTACAAGGAAATGGGTATCGGTATCATCATTCCTCAAAATCTTCTTGACAA 924

```



```
Db 2021 TATGATGTATGAAGTCCGGTATTGTTGAGACTTTTGAAGAGGTCTCAGAAATGTC 2080
Qy 925 TTCGTTCCCTCTTTTTCAGGTT- ACTATTGATCCAGCTTCACACCCACAGCTCCATGTCCT 983
Db 2081 TTTGAACCACTGTTTGAAGTTTCAAAAGGATCTCGTACACATCTCTTAAACTCCACGTGTT 2140
Qy 984 CCTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACACGAAAGGGTCC 1043
Db 2141 TTTGCAACGTGTTATCGGGTTTGACTCTGTGACGACGAAAGTAAACCTGAAACGCCGTAC 2200
Qy 1044 AACAAAGCAGTCCGCCACACCTGACAGTGGACCAATGTTCAACCCCTGCATTTTCATA 1103
Db 2201 CTTTCGCAATTCCTTATCCGAAACATTTGGGATATTAATTTAAATCCTCCATATAGTTA 2260
Qy 1104 TTATGCGTACTACTGTCTACTTAACCTATTTCACCCCTAAACAAAGCTGCGTGAAGGG 1163
Db 2261 TTGGTTATATTACATGATGCAACATGACTAGCTTGAATAGCTGGAGAAAAATACGTGG 2320
Qy 1164 AATGACCACATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGTTGATCACTTGGC 1223
Db 2321 TTTTAAACACCTTTGTCTTCGTCCTCATTTGGTGAAGCTGGAGATACGGATCAITTTGGC 2380
Qy 1224 AGCGACATTTCTCTCTGTCAGAACATATCACATGGAATTAATCTAAGGAAGTCTCCTGT 1283
Db 2381 CAGTGTCTTTCTTATTAAGTCAAGTATTAACCATGGTATTTTATTACGAAAGTTCCTTT 2440
Qy 1284 GCTTCAGTACTTGTACTATCTTTGGTCAGATTTGGTCTGGCGATGTCCCCATTGAGCAACAA 1343
Db 2441 TCTTCAGTATCTTTGGTACCTTGATCAGATCCCTATCGCAATGTCTCCATTATCAACAA 2500
Qy 1344 CTCCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAAGTTCCTCCAAAGAGGTCTGAA 1403
Db 2501 TGCACCTGTTTCTTGCAATATGATAAGAAATCCATTTTGAAGTACTTCAAAACGAGGTCTTAA 2560
Qy 1404 TGTCTCATTTATCTACGGATGACCCCTTTGCAAAATTCACCTGACAAAGAACCATTTGGTGA 1463
Db 2561 TGTAGTCTTTCCACTGATGATCCCTTTACAAITTGCAITTTACCCGAGAGCCTCTGATTGA 2620
Qy 1464 AGAATACAGCATTTGCTGCTGCTGTGGAAGCTCAGTTCTTGTGATTTATGCGAAATTGC 1523
Db 2621 GGAATATGCTGTAGCGCACAAATTTATAAGCTCTCTGCTGTCGATATGTGCGAATTGCG 2680
Qy 1524 GAGGAACCTCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGCGGCACTGGATTGGT 1581
Db 2681 TAGAAAACAGTGTTTTACAATCCGGTTTCGAACGTCAGCTTAAAGAACGTTGGTTGGGT 2738
```

Search completed: December 15, 2003, 20:04:24  
Job time : 6853 secs





PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;  
XX WPI; 2001-159866/16.  
DR P-PSDB; AAB31948.  
XX  
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used  
PT e.g. for designing or identifying herbicides that inhibit the enzyme  
PT activities, and as probes for genetic or physical mapping  
XX  
PS Claim 2; Page 42; 72pp; English.  
XX  
CC The present sequence encodes an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).  
XX  
SQ Sequence 1910 BP; 596 A; 391 C; 383 G; 540 T; 0 other;  
  
Query Match 100.0%; Score 1910; DB 22; Length 1910;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 CCACGGCTCGGGTGTCTCCATCGGAGAGAGGTCATAAATCACCCCTGTACTCCAAA 60  
DB 1 CCACGGCTCGGGTGTCTCCATCGGAGAGAGGTCATAAATCACCCCTGTACTCCAAA 60  
  
QY 61 CCTAACCCCAACCGGTTCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120  
DB 61 CCTAACCCCAACCGGTTCACTTATGTGCTGAAACCAAGTCAGAGCATGTTTCCAAACT 120  
  
QY 121 GTTGATGGCGTTATCCATGTTTATGGGATAAAGATTGTACGAGAGCAATTTATCTGTG 180  
DB 121 GTTGATGGCGTTATCCATGTTTATGGGATAAAGATTGTACGAGAGCAATTTATCTGTG 180  
  
QY 181 GCTGATGCTACAACTTCTTCACTGACTTGCATTATATCTCCGAGTAACGGCTGCAGGG 240  
DB 181 GCTGATGCTACAACTTCTTCACTGACTTGCATTATATCTCCGAGTAACGGCTGCAGGG 240  
  
QY 241 AACACAGAAGACTGTCTGCCATAATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCAT 300  
DB 241 AACACAGAAGACTGTCTGCCATAATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCAT 300  
  
QY 301 CTGATGTTAAATGGGATAGGAATTTCTTGGCCAGAGAGACTGCCCCACATCGTATTTT 360  
DB 301 CTGATGTTAAATGGGATAGGAATTTCTTGGCCAGAGAGACTGCCCCACATCGTATTTT 360  
  
QY 361 TACAATGTCAGGAAGGTTGACACTCATGTTTCATCTTACATTCAGCATGTAATCAAAACAT 420  
DB 361 TACAATGTCAGGAAGGTTGACACTCATGTTTCATCTTACATTCAGCATGTAATCAAAACAT 420  
  
QY 421 CTGTTGAGGTTTCAATAATCAAACTAAGAAAAAGAACCTGATGAGGTGGTCAITTTTCA 480  
DB 421 CTGTTGAGGTTTCAATAATCAAACTAAGAAAAAGAACCTGATGAGGTGGTCAITTTTCA 480  
  
QY 481 GATGGTACTTATATGACTTTAAAGAGGTTTTTGGAGAGCTTGAACCTGAGGTATGAT 540  
DB 481 GATGGTACTTATATGACTTTAAAGAGGTTTTTGGAGAGCTTGAACCTGAGGTATGAT 540  
  
QY 541 CTGAATGTTGATTTGCTAGATGTCCATGTCAGACAAAGACATTTTCATCGTTTTCACAAA 600

541 CTGAATGTTGATTTGCTAGATGTCCATGTCAGACAAAAGCACATTTTCATCGTTTTCACAAA 600  
601 TTCAATCTAAATACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAACAA 660  
601 TTCAATCTAAATACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAACAA 660  
661 GATATCTTATTTCAAGGCGGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTT 720  
661 GATATCTTATTTCAAGGCGGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTT 720  
721 TCTGCTAGCAATATACAGATGCGAGATATAGGATTTCAATCTACGGAAGGAAACAGAGT 780  
721 TCTGCTAGCAATATACAGATGCGAGATATAGGATTTCAATCTACGGAAGGAAACAGAGT 780  
781 GAATGGGACCAACTTGAAGTTGGATGTAACATGTAATGTCACAGTGGAAATGTTCTC 840  
781 GAATGGGACCAACTTGAAGTTGGATGTAACATGTAATGTCACAGTGGAAATGTTCTC 840  
841 TGGCTGGTTTCAGATTCACAGCTTATATATGTAACAGGAAATGGGTATCGTTACATCA 900  
841 TGGCTGGTTTCAGATTCACAGCTTATATATGTAACAGGAAATGGGTATCGTTACATCA 900  
901 TTCCAAAATCTTTTGACAAACATTTTCCTCTTTTGGAGTTTACTTATGATCCAGCT 960  
901 TTCCAAAATCTTTTGACAAACATTTTCCTCTTTTGGAGTTTACTTATGATCCAGCT 960  
961 TCACACCCACAGCTCCATGTTCTTCTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGAT 1020  
961 TCACACCCACAGCTCCATGTTCTTCTGAAGCAGGTTGTAGGTTGGACCTGGTTGATGAT 1020  
1021 GAAAGTAAACCAAGGCGTCCAAACAAAGCACATGCCACACCTGAAACAGTGGACCAAT 1080  
1021 GAAAGTAAACCAAGGCGTCCAAACAAAGCACATGCCACACCTGAAACAGTGGACCAAT 1080  
1081 GTGTTCAACCCCTGCATTTTTCATATATGCTACTACTGCTATGCTAACTTATTCACCCCTA 1140  
1081 GTGTTCAACCCCTGCATTTTTCATATATGCTACTACTGCTATGCTAACTTATTCACCCCTA 1140  
1141 AACAGCTGCGTGAGTCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200  
1141 AACAGCTGCGTGAGTCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200  
1201 GCTGAGAGTGTGATGACTTGGCAGCAGCATTTCTTCTGTCACAAACATATCAGATGA 1260  
1201 GCTGAGAGTGTGATGACTTGGCAGCAGCATTTCTTCTGTCACAAACATATCAGATGA 1260  
1261 ATTAATCTAAGGAAGTCTCTGTGTTTCTGATCTAGTACTATCTTGGTTCAGATTTGCTG 1320  
1261 ATTAATCTAAGGAAGTCTCTGTGTTTCTGATCTAGTACTATCTTGGTTCAGATTTGCTG 1320  
1321 GCGATGTCCTCAATTGAGCAACATCTTATTTCTTGACTATCATCGAACCCCTTTTCCA 1380  
1321 GCGATGTCCTCAATTGAGCAACATCTTATTTCTTGACTATCATCGAACCCCTTTTCCA 1380  
1381 AGTCTTCTTCAACAGAGTCTGAATGTCATTTATCTACGATGACCCCTTGGCAATTCAC 1440  
1381 AGTCTTCTTCAACAGAGTCTGAATGTCATTTATCTACGATGACCCCTTGGCAATTCAC 1440  
1441 CTGCAAAAGAACCCATTTGGTGGGAAGATAACAGCATTTGCTGCTGCTGGAAGTTCAGT 1500  
1441 CTGCAAAAGAACCCATTTGGTGGGAAGATAACAGCATTTGCTGCTGCTGGAAGTTCAGT 1500  
1501 TCTTGTGATTTATGCGAAATTCGAGGAACCTCTGTTTACCATCTGGGTTTTCACATGCT 1560  
1501 TCTTGTGATTTATGCGAAATTCGAGGAACCTCTGTTTACCATCTGGGTTTTCACATGCT 1560  
1561 CTCAGGGCCACTGGAATTTGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAATGATATT 1620  
1561 CTCAGGGCCACTGGAATTTGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAATGATATT 1620  
1621 CACAGAACCAATGTATCCGCACATCAGGGTTCAATTTAGAGAGATGATCTGGAGAAATGAA 1680

Db 1621 CACAGAACCAATGTACCGCACATCAGGGTTCATTTAGAGAGATGATCTGGAGAAATGAA 1680  
 QY 1681 ATGAAACTAGTGTACTGTGCAATGAGATCTTAATACCGAGAGCTGGACCTGTGAAT 1740  
 Db 1681 ATGAAACTAGTGTACTGTGCAATGAGATCTTAATACCGAGAGCTGGACCTGTGAAT 1740  
 QY 1741 GTCCAGCCTCGTGTATACAGAGAGCTGGTGTGTAGCTGTATGGGAATTTACTTCAT 1800  
 Db 1741 GTCCAGCCTCGTGTATACAGAGAGCTGGTGTGTAGCTGTATGGGAATTTACTTCAT 1800  
 QY 1801 GTTTGGTATGCTTTCTTCTATCTATGCGCAATTCAACTTCGAACTTCAAAAAA 1860  
 Db 1801 GTTTGGTATGCTTTCTTCTATCTATGCGCAATTCAACTTCGAACTTCAAAAAA 1860  
 QY 1861 AAAAAA 1910  
 Db 1861 AAAAAA 1910

## RESULT 2

AAF25494

ID AAF25494 standard; DNA; 2573 BP.

XX

AC AAF25494;

XX 15-MAY-2001 (first entry)

XX

DE Nucleotide sequence of a soybean AMP deaminase enzyme.

XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;

KW inosine; ss.

XX Glycine max.

OS

XX Key

FT 652..2463

FT CDS

FT /\*tag= a

FT /product= "AMP deaminase"

FT /transl\_except= "(pos: 694..696, aa: Xaa)"

FT /note= "Xaa is an unknown residue"

XX WO200109305-A2.

XX

PD 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US21009.

XX

XX 30-JUL-1999; 99US-0146473.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.

XX

XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

PI

XX WPI: 2001-159866/16.

XX P-PSDB; AAB31954.

DR

XX New polynucleotides encoding AMP deaminase or adenosine deaminase used

PT e.g. for designing or identifying herbicides that inhibit the enzyme

PT activities, and as probes for genetic or physical mapping

XX

PS Claim 2; Page 56-57; 72pp; English.

XX

CC The present sequence encodes an AMP deaminase. The specification also  
 CC describes adenosine deaminase. These enzymes convert adenosine to  
 CC inosine. Mutations in these genes cause disruptions in then salvage and  
 CC catabolism of adenosine and AMP. In humans, this may lead to death of  
 CC white blood cells, which causes severe immunodeficiencies. The AMP  
 CC deaminase and adenosine deaminase may be used to prepare antibodies  
 CC to these proteins, and to design or identify herbicides that inhibit  
 CC their enzyme activities. The polynucleotides are used as probes for  
 CC genetically and physically mapping genes that they compose, and as  
 CC markers for traits linked to those genes, where such information may be

CC used in plant breeding to develop lines with desired phenotypes. The  
 CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
 CC homologous proteins from the same or other plant species, and in  
 CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
 CC acid fragments encoding homologous genes from DNA or RNA. These may  
 CC also be used to create transgenic plants in which the polypeptides are  
 CC overexpressed or suppressed, and as probes in direct fluorescent in  
 CC situ hybridisation (FISH).

XX Sequence 2573 BP; 716 A; 523 C; 588 G; 694 T; 52 other;

Query Match 96.1%; Score 1836; DB 22; Length 2573;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1836; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GGTGCTCCATGGGAGAGAGGTGATAAATGACCCCTGTACTCCAAACCTTAACCCCA 71

Db 738 GGTGCTCCATGGGAGAGAGGTGATAAATGACCCCTGTACTCCAAACCTTAACCCCA 797

QY 72 CCGTTCACTTATGTGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGTGATGGCT 131

Db 798 CCGTTCACTTATGTGCTGAACCAAGTCAGAGCATGTTTCCAAACCTGTGATGGCT 857

QY 132 TATCCATGTTTATGGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGATGCTAC 191

Db 858 TATCCATGTTTATGGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGATGCTAC 917

QY 192 AACCTTCTCACTGACTTGCATTATATCTCCGAGTACGGCTGAGGAGACACAGAAGAC 251

Db 918 AACCTTCTCACTGACTTGCATTATATCTCCGAGTACGGCTGAGGAGACACAGAAGAC 977

QY 252 TGTCTGCCAATATCGGTTAAATCTTTTGAGCATAAGTTTAAATTCATCTGATGTTAAA 311

Db 978 TGTCTGCCAATATCGGTTAAATCTTTTGAGCATAAGTTTAAATTCATCTGATGTTAAA 1037

QY 312 TGGGATAGGAAATTTCTTGCCGAAGACATGCCCCACATCGTGAATTTTCAATGTGAG 371

Db 1038 TGGGATAGGAAATTTCTTGCCGAAGACATGCCCCACATCGTGAATTTTCAATGTGAG 1097

QY 372 GAAGTTGACACTCATGTTTCATCTCAGCATGATGAATCAAAAACATCTGTTGAGGTT 431

Db 1098 GAAGTTGACACTCATGTTTCATCTCAGCATGATGAATCAAAAACATCTGTTGAGGTT 1157

QY 432 CATAAATCCAACTAAGAAAAGAACCTGTAGTGGTGTGCTCAATTTTCAGAGATGTTACTTA 491

Db 1158 CATAAATCCAACTAAGAAAAGAACCTGTAGTGGTGTGCTCAATTTTCAGAGATGTTACTTA 1217

QY 492 TATGACTTTAAAGAGAGGTTTTTGAGAGCTTGGACTTAACCTGGGTATGATCTGAATGTTGA 551

Db 1218 TATGACTTTAAAGAGAGGTTTTTGAGAGCTTGGACTTAACCTGGGTATGATCTGAATGTTGA 1277

QY 552 TTTGCTAGATGCTCCATGCAGACAAAGACATTTTCATCGTTTTCAGCAAAATTCATCTTAA 611

Db 1278 TTTGCTAGATGCTCCATGCAGACAAAGACATTTTCATCGTTTTCAGCAAAATTCATCTTAA 1337

QY 612 ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAAGATAATCTTAT 671

Db 1338 ATACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAAGATAATCTTAT 1397

QY 672 TCAAGGGCGTTTCTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGTAGCAA 731

Db 1398 TCAAGGGCGTTTCTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGTAGCAA 1457

QY 732 ATATCAGATGCGCAATATAGGTTTCAATCTACGGAGGAAACAGATGATGATGGACCA 791

Db 1458 ATATCAGATGCGCAATATAGGTTTCAATCTACGGAGGAAACAGATGATGATGGACCA 1517

QY 792 ACTTGCAGTTGGATAGTGAACCAATGAATTCAGAGTGAATTTCTCTGGCTGGTTCA 851

Db 1518 ACTTGCAGTTGGATAGTGAACCAATGAATTCAGAGTGAATTTCTCTGGCTGGTTCA 1577

QY 852 GATTCCACGCTTATATATGTGTAAGGAAATGGGTATCGTTTACATCATTTCCAAAATCT 911

```
Db 1578 GATTCACGCTTATATATGCTACAAGGAATGGGTATCGTTATCATCATTCACAAATCT 1637
Qy 912 TCTTGACAACTTTTCCTCTCTTTTCAGGTTACTATTGATCCAGCTTCACACCCACA 971
Db 1638 TCTTGACAACTTTTCCTCTCTTTTCAGGTTACTATTGATCCAGCTTCACACCCACA 1697
Qy 972 GCTCCATGCTTCCTGGAACAGGTTGTAGGGTTGACCTGGTTGATGATGAAGTAACCC 1031
Db 1698 GCTCCATGCTTCCTGGAACAGGTTGTAGGGTTGACCTGGTTGATGATGAAGTAACCC 1757
Qy 1032 AGAAGGCGTCCAACAAAGCAGATGCCACACCTGAACAGTGGACCAATGTGTTCAACCC 1091
Db 1758 AGAAGGCGTCCAACAAAGCAGATGCCACACCTGAACAGTGGACCAATGTGTTCAACCC 1817
Qy 1092 TGCATTTTCATATTATGCGTACTACTGCTATGCTATGCTATGCTATGCTATGCTATGCT 1151
Db 1818 TGCATTTTCATATTATGCGTACTACTGCTATGCTATGCTATGCTATGCTATGCTATGCT 1877
Qy 1152 TGAGTCAAAAGGAATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGT 1211
Db 1878 TGAGTCAAAAGGAATGACCACTATCAAAATTCGGTCCACATGCTGGAGAGGCTGGAGATGT 1937
Qy 1212 TGATCATTTGGCAGCAGCAATTTCTCTCTGTCAACAATATCACATGGAATTAATCTAAG 1271
Db 1938 TGATCATTTGGCAGCAGCAATTTCTCTCTGTCAACAATATCACATGGAATTAATCTAAG 1997
Qy 1272 GAAGTCTCTGTGCTTCACTACTTGTACTATCTTGTGTCAGATGCTGCGCATGCTCCCC 1331
Db 1998 GAAGTCTCTGTGCTTCACTACTTGTACTATCTTGTGTCAGATGCTGCGCATGCTCCCC 2057
Qy 1332 ATTGAGCAACAACCTCTTATTTCTGACTATATCGCAACCCCTTTTCCAAAGCTTCTTCCA 1391
Db 2058 ATTGAGCAACAACCTCTTATTTCTGACTATATCGCAACCCCTTTTCCAAAGCTTCTTCCA 2117
Qy 1392 ACGAGTCTGAATGTCTCAATATCTACGGATGACCCCTTTTGCAAATTCACCTGACAAAGA 1451
Db 2118 ACGAGTCTGAATGTCTCAATATCTACGGATGACCCCTTTTGCAAATTCACCTGACAAAGA 2177
Qy 1452 ACCATTGGTGAAGATACAGATTCGCTTCTGCTGCTGGAAGCTCAGTTCTTGATTT 1511
Db 2178 ACCATTGGTGAAGATACAGATTCGCTTCTGCTGCTGGAAGCTCAGTTCTTGATTT 2237
Qy 1512 ATGCGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCA 1571
Db 2238 ATGCGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCA 2297
Qy 1572 CTGGATTGTAAGAACTACTTCAAAAGAGACCTGCTGGAATGATATTCACAGAACCAA 1631
Db 2298 CTGGATTGTAAGAACTACTTCAAAAGAGACCTGCTGGAATGATATTCACAGAACCAA 2357
Qy 1632 TGTACCGCATCAGGGTTCMAATTTAGAGAGATGATCTGGAGAAATGAATGAACACTAGT 1691
Db 2358 TGTACCGCATCAGGGTTCMAATTTAGAGAGATGATCTGGAGAAATGAATGAACACTAGT 2417
Qy 1692 GTACTCTGCAATGAGATCTTAATACACAGCAGCTGGACCTGAAGATGTCAGGCTCG 1751
Db 2418 GTACTCTGCAATGAGATCTTAATACACAGCAGCTGGACCTGAAGATGTCAGGCTCG 2477
Qy 1752 TGTATACAGACAGTGGCTGTAGCTGCTATGGAATPATCTTCACTTTTGTGATG 1811
Db 2478 TGTATACAGACAGTGGCTGTAGCTGCTATGGAATPATCTTCACTTTTGTGATG 2537
Qy 1812 CTTTCTCTATCTATGCAAAATTCACCTTGCACCTC 1847
Db 2538 CTTTCTCTATCTATGCAAAATTCACCTTGCACCTC 2573
```

RESULT 3

AAF25497

ID AAF25497 standard; DNA; 1988 BP.

XX

AAF25497;

XX

```
DT 15-MAY-2001 (first entry)
XX Nucleotide sequence of a soybean adenosine deaminase enzyme.
DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
XX inosine; ss.
KW Glycine max.
OS
XX Key Location/Qualifiers
FH 661..1697
CDS /*tag= a
FT /product= "adenosine deaminase"
FT /transl_except= "(pos: 670..671, aa: Phe)"
XX
XX WO200109305-A2.
XX
XX 08-FEB-2001.
XX
XX 28-JUL-2000; 2000WO-US21009.
XX
XX 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
PA (PION-) PIONEER HI-BRED INT INC.
XX
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
DR P-PSDB; AAB31957.
XX
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX
XX Claim 2; Page 65-66; 72pp; English.
XX
XX The present sequence encodes an adenosine deaminase. The specification
CC also describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX
XX Sequence 1988 BP; 574 A; 420 C; 429 G; 565 T; 0 other;
SQ
Query Match 64.8%; Score 1238; DB 22; Length 1988;
Best Local Similarity 86.4%; Pred. No. 2.3e-218;
Matches 1402; Conservative 0; Mismatches 215; Indels 5; Gaps 3;
Qy 122 TTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGACGAGAGCATTTATCTCTG-TG 180
Db 83 TGGCTGTTGTTCCCAAGTCTTTGTAGAATAAAGTCGATCAAAAGGATTATCTCTGTTT 142
Qy 181 GCTGATGCTACA-ACCTTCTTCACCTGACTTGATTTATATATTTCTCCGAGTAACCGCTG 239
Db 143 GCTGATGACGAGACCTTTTCCCGACTTACATTTATGTTCTCCGGTGACTCCGCGG 202
Qy 240 GAACACAGAACTGTCTGCCATAATCGGTTAAATCTTCTTGAGCATAAGTTTAAATCCA 299
Db 203 GAACACAGAACTGTCTGCCATAACCGATTGAATCTTCTAGAACATCAAGTTCAATTTCA 262
```





CC inosine. Mutations in these genes cause disruptions in then salvage and  
 CC catabolism of adenosine and AMP. In humans, this may lead to death of  
 CC white blood cells, which causes severe immunodeficiencies. The AMP  
 CC deaminase and adenosine deaminase may be used to prepare antibodies  
 CC to these proteins, and to design or identify herbicides that inhibit  
 CC their enzyme activities. The polynucleotides are used as probes for  
 CC genetically and physically mapping genes that they compose, and as  
 CC markers for traits linked to those genes, where such information may be  
 CC used in plant breeding to develop lines with desired phenotypes. The  
 CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
 CC homologous proteins from the same or other plant species, and in  
 CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
 CC acid fragments encoding homologous genes from DNA or RNA. These may  
 CC also be used to create transgenic plants in which the polypeptides are  
 CC overexpressed or suppressed, and as probes in direct fluorescent in  
 CC situ hybridisation (FISH).

XX  
 SQ Sequence 2782 BP; 798 A; 640 C; 591 G; 753 T; 0 other;

Query Match 57.7%; Score 1101.4; DB 22; Length 2782;  
 Best Local Similarity 78.5%; Pred. No. 2.9e-193;  
 Matches 1318; Conservative 0; Mismatches 361; Indels 0; Gaps 0;

QY	13	GTGCTCCATGGGAGAGGAGTCAATAATGACCCCTGACTCCAAACCTTAACCCCAAC	72
DB	768	GTGCTCCATGGGAGAGGAGTCAATAATGACCCCTGACTCCAAACCTTAACCTTAAC	827
QY	73	CGGTTCACCTATGTGCTGACCAAGTCAGACATGTTTCCAAACTGTGATGCGGTT	132
DB	828	CGTTTCTATTACGAGCAGCAGACTAAACCTGAACATCATTTTGAATGGTTGATGGT	887
QY	133	ATCCATGTTTATGCGGATTAAGATGTCGGAGAGCATTTATCCTGCTGGCTGATGCTACA	192
DB	888	ATTTCATGTTATACCCCAATTAAGACGCTTAAGAAAGAAATCTATCCTGCTGATGCTACT	947
QY	193	ACCTTCTTCACTGACCTGATTTATTTCTCCAGTAACCGCTGACGAGGAAACAAAGAACT	252
DB	948	ACCTTTTCTGATGATGACATATATCCTTCGTGTTGCTGCTGGGATATTCGAAT	1007
QY	253	GTCTGCCAATAACGGTTAAATCTTCTTGAGCATAAAGTTAAATTCATCTGATGTTAAAT	312
DB	1008	GTATGTTATAAAGCTTTAAATCTTCTAGAACAGAAATTCATCTTCAATTTGATGGTCAAT	1067
QY	313	GGCGATAGGGAATTTCTTGCCAGAGAGCTGCCCAACATGCTGATTTTACATATGTCAGS	372
DB	1068	GCGGATAGAGAACTACTTGCTCAGAAAGCTGACCCCTCGGACTTCTACAAATGTCAGG	1127
QY	373	AAGGTTGACACTCATGTTTCATCATTTCCAGCATGATGAATCAAAACATCTGTTGAGGTTTC	432
DB	1128	AAGGTTGATACCTCATGTTTCATCACTCTGCATGATGAATCAGAGCAATTTGTTGAGATT	1187
QY	433	ATAAATCCAACTAAGAAAGAACCTGATGAGGTGGTCATTTTCAGAGATGTCATTAT	492
DB	1188	ATCAAGTCCAAAGTTGAGGAAGAACCTGACGAGGTTGTGATTTTATAGAGATGTCCTAT	1247
QY	493	ATGACTTTAAAGAGGTTTTTGAGAGCTTGAGCTTAACCTGGGTATCATCTGATGTTGAT	552
DB	1248	TTGACTTCTTAAGAGGTTTTTGAGAGTTTGGACTTGACTGGTTTATGACCTCAATGTTGAT	1307
QY	553	TTCTAGATGTCATGACAGCAAAAGCACATTTTCATCGTTTTCAGCAAAATTCATCTAAAAA	612
DB	1308	CTCTAGATGTCATGCCGATAAAGTACATTCCTCCATCGCTTGACAAAGTTCAATTTGAAG	1367
QY	613	TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAACAAGATATCTTATT	672
DB	1368	TATAATCTTGTGGCCAAATCCCGGCTGAGGAGATCTTTCTTAAACAGGACACCTTATT	1427
QY	673	CAAGGCGGTTTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAA	732
DB	1428	CAAGGCGGATTTCTTCTGTAATTGACAAAGAAAGTATTTTCTGATCTTGAAGCAAGTAA	1487
QY	733	TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA	792

DB	1488	TATCAGATGGCTGAGTATAGAAATATCTATCTATGGGAGAGAAAGAAAGTGGGATCAG	1547
QY	793	CTTGCAAGCTTGATAGTCAACAAATGGAATTTGCACAGTGGAAATGTTGTTCTGGCTGTTCA	852
DB	1548	ATGGCAAGCTTGATAGTGAATAATGTAACACGAGAAATGTTGTTGGTTAATTCAG	1607
QY	853	ATTCCACGCTTATATAATGTTGTACAAGGAAATGGGTATCGTTTACATCATTTCCAAATCTT	912
DB	1608	ATTCTCGATATACAATGTATACAGGAGATGGGAACAATCAATTCITTTCCAGAACCTC	1667
QY	913	CTTGACAACATTTTCGTTCTCTTTTGGAGTTACTATTGATCCAGCTTTCACACCCACAG	972
DB	1668	CTTGACAATATTTTCTGCTCTTTTGAAGTAACTGTTGATCTCTGCTTTCACATCTCAG	1727
QY	973	CTCCATGTTCTCCCTGACAGCAGTTGTAGGTTTGGACCTGTTGATGATGAAGTAAACCA	1032
DB	1728	CTCCAATGTTTCTTGCAACAGCTGTTGGGCTGGATTTAGTGATGATGAAGCAACCA	1787
QY	1033	GAAAGGCGTCCAAACAAAGCACATGCCACACCTCGAACAGTGGACCAATGTGTTCAACCTC	1092
DB	1788	GAGAGACGCCCAACAACAACACATGCTACACCTGAGCAATGACTAATGTTTCAATCCA	1847
QY	1093	GCATTTTCAATATATGCGTACTACTGCTATGCTAACTTATTCACCTTAAACAAGCTGCT	1152
DB	1848	GCATATGCATATATGCTGCTACTTATGTTTATGCTAACTTGTACACGCTGAACAAGCTCGT	1907
QY	1153	GAGTCAAAGGGAATGACCACTATCAAAATCCGTCACACATGCTGGAGAGCTGGAGATGTT	1212
DB	1908	GAGTCCAAAGGATGACCAACAACTTCTGTCACACATGCTGGGAGGCTGGAGATAT	1967
QY	1213	GATCACTTTGGGACGACACATTTCTCTCTGTGCACACATATCACATGGAATTAATCTAAGG	1272
DB	1968	GATCATCTTGTGCGAGCATTTCTTACTTCTCATATATTTGCTCAGGGGTAAATTTAAG	2027
QY	1273	AAGTCTCTGCTTTCAGTACTGTTGATCTTGTGTCAGATGTTGCTGGCGATGTCCTCA	1332
DB	2028	AAGTCCCTGCTCCTCCAGTATCTGTATTACCTAGCTCAGATGTTGCTTGGCCTCTCTCT	2087
QY	1333	TTGAGCAACAACCTCTTATTTCTTGACTATCATCCCAACCTTTTCCACGTTCTTCCAA	1392
DB	2088	TTGAGCAACAACCTCTTATTTGATTTATGATTTATCACCGAAACCTTTTCCCAACAT	2147
QY	1393	CGAGTCTGAAATGCTCTCTATCTACGATGACCCCTTTGCAAAATTCACCTGACAAAAGAA	1452
DB	2148	AGAGGCTTTACGTTTCTCTATCAACCGATGACCTTTGCAAAATTCACCTGACAAAAGAA	2207
QY	1453	CCATTGGTGGAGAAATACAGCATGCTGCTGCTGCTGGAGAGCTCAGTTCTTTGATTTA	1512
DB	2208	CCTTTGGTTGAAGAAATATAGCATCGCTGCTGCTGCTGGAGAGCTAAGTTTCATGCGACCTA	2267
QY	1513	TGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGGCAC	1572
DB	2268	TGTGAAATTCGTAGGAAATCTGTGTACCAATCTGGTCTCTCTCATAGGCTCAAGTCACAC	2327
QY	1573	TGCAATTTGTAAGAACTACTTCAAAAGAGGAGCCTGCTGGAAATGATATTTCACAGAACCAAT	1632
DB	2328	TGCAATTTGGAGAACTACTTCAAAAGAGGAGCCTGCTGGAAATGATATTTCACAGAACCAAT	2387
QY	1633	GTACCGCATCATCAGGTTTCAATTTTAGAGAGATGATCTGGAGAAATGAAATGAAATCTAGT	1691
DB	2388	GTTCTCTCATCATCAGGATTTGAATTCGACACACTATTGGAAGAGAAATGAGAGCTAAT	2446

RESULT 5  
 AAL45016  
 ID AAL45016 standard; DNA; 2520 BP.  
 XX  
 AC AAL45016;  
 XX  
 DT 16-MAY-2002 (first entry)  
 XX  
 DE A thaliana AMP deaminase coding sequence.  
 XX

KW AMP deaminase; adenosine monophosphate; transgenic plant;  
 KW herbicide resistance; herbicide; inhibitor; gene; ds.  
 XX Arabidopsis thaliana.

XX Key Location/Qualifiers  
 XX 1..2520  
 FT /\*tag= a  
 FT /product= "AMP deaminase"  
 XX

PN WO200206319-A2.  
 XX  
 XX  
 PD 24-JAN-2002.  
 XX  
 XX  
 PF 06-JUL-2001; 2001WO-BP07767.  
 XX  
 XX  
 PR 17-JUL-2000; 2000DE-1035084.  
 XX  
 XX (AVET.) AVENTIS CROPS SCIENCE GMBH.  
 PA  
 XX Schulz A, Streber W, Hanke C, Schmidt F, Schubel A;  
 XX WPI; 2002-195802/25.  
 DR P-PSDB; AA016943.  
 DR

XX New nucleic acid for plant adenosine monophosphate deaminase, useful in  
 PT screening, for herbicides and preparing herbicide-resistant plants -  
 XX  
 XX Claim 1; Page 41-44; 51pp; German.

CC The present invention provides the protein and coding sequences of the  
 CC Arabidopsis thaliana adenosine monophosphate (AMP) deaminase. The coding  
 CC sequence can be used to transform prokaryotic or eukaryotic cells,  
 CC especially to produce transgenic plants (e.g. barley, rice, soya etc.)  
 CC resistant to herbicidal inhibitors of AMP-deaminase, for recombinant  
 CC production of proteins with AMP-deaminase activity, and to identify  
 CC related genes in other organisms. The protein can be used for  
 CC identification and biochemical/structural characterisation of new  
 CC AMP-deaminase inhibitors and potential herbicides. The present sequence  
 CC is the coding sequence of the invention.

XX Sequence 2520 BP; 699 A; 541 C; 572 G; 708 T; 0 other;

Query Match 54.6%; Score 1043.4; DB 24; Length 2520;

Best Local Similarity 76.0%; Pred. No. 1.3e-182;

Matches 1287; Conservative 0; Mismatches 406; Indels 0; Gaps 0;

QY	13	GTGCTCCATGGGAGAGGCTATAATGACCCCTGTACTCCAAACCTTAACCCCAAC	72
DB	799	GTGACACATGGGAAAGAGTCAATCTGATCTCTAGTCTCAAAAGCCTTAATACAGAG	858
QY	73	CCGTTTCACTATGTGCTGAACCAAGTCAGAGCATGTTTCCAAACTGTTGATGGCGTT	132
DB	859	CAATTGACACATATCTCAGGAAATCTGATCATGTTTGGATGCAAGATGGGTT	918
QY	133	ATCCAGTTTATCGGATAAAGATTGTACGGAGACATTTATCCGTGGCTGATGCTACA	192
DB	919	GTCCACGTTTGGCAATAAAGATGCAAAAGAAAGTCTCTCCGGTAGCTGATGCCACA	978
QY	193	ACCTTCTTCACGTGCTGATTAATTTCTCCAGTAACCGCTGCGAGGAAACACAAACT	252
DB	979	GGGTTTTTCACTGATGTCATCACGTACTCTCAAGTCATAGCTGCGAGAAACATCCGGACT	1038
QY	253	GTCTGCCAATATCGGTTAAATCTCTTGAGCATTAAGTTTAAATTCATCTGATGTTAAT	312
DB	1039	TTTGCCACCGTCGACTAGTGTCTCTAGACAGAAATTAATCTCAATTTGATGTTAAT	1098
QY	313	CGGATAGGAAATTTCTTGCCAGAGACTGCCCAACATCGTGAATTTTACAATGTCAGG	372
DB	1099	CGGGATAAAGAAATTTCTTGCTCAAAAAGTGCACACATCGTGAATTTTATAACGTAGG	1158
QY	373	AGGTTGACACTCATGTTCAATTCAGATGTCATGAATCAAAAACATCTGTTGAGGTTT	432

DB	1159	AAAGTCGACACTCATGTGCATCATTCAGCTTGCATGAACCAAGAAACACCTTTTAAGGTTT	1218
QY	433	ATAAAATCCAACTAAGAAAGAAACCTGATGAGGTGGTCATTTTCAGAGATGGTACTTAT	492
DB	1219	ATTAAGTCAAGCTCCGGAAGAACCCGATGAGGTAGTAATAATTCGAGATGGACATAT	1278
QY	493	ATGACTTTAAAGAGGTTTTTGGAGAGCTTGGACTTAACTGGGTATGATCTGTAATTTGAT	552
DB	1279	TTGACCTTTGAGAGAGTTTTTTGAGAGCCTGGATCTGACTGGGTATGACTGAACCTCGAC	1338
QY	553	TTGCTAGATGCTCATGACAGAAAGACATTTCTATCGTTTTCACAAATTCATCTAAAA	612
DB	1339	CTTTTGGATGTTTCATGACAGAAAGTACCTTTTCATCGTTTTCATGATGTTCAACCTAAG	1398
QY	613	TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCCCTCAAAACAGATAATCTTAT	672
DB	1399	TATAACCTTTGTTGTTCAAGTAGGCTTAGGGAGATTTTCCCTTAAACAGATAATCTCATC	1458
QY	673	CAAGCCCTTTTCTGCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAA	732
DB	1459	CAAGTCGATTTCTTGGTGAGATAACAAAGCAAGTCTTCTCTGACCTTTGAAGCTAGTAA	1518
QY	733	TATCAGATGGCAGATATAGGATTTCAATCTACGGAAGGAACAGAGTGAATGGACCAA	792
DB	1519	TATCAGATGGCTGAATACAGAAATATCTATATATGSCAGAAATGAGCGAGTGGACCAA	1578
QY	793	CTTCAAGTTGGATAGTGAACAATGCAATGCAAGTGGAAATGTTGCTGGCTGGTTCCAG	852
DB	1579	CTCGTAGTTGGATTGTGAACAATGATCTATACAGTGAATGTTGCTGTTAATTCAG	1638
QY	853	ATTCACGCTTATATAATGTGTACAAGAAATGGGTATCGTTTACATCAATCCAAATCTT	912
DB	1639	CTCCACGCTTGTACAACATTTACAAGGACATGGTATTTGTGACATCGTTCAGATAATC	1698
QY	913	CTTGACACATTTTGGTTCCTCTTTTGGAGTTACTATTGATCCAGTTTACACCCACAG	972
DB	1699	CTGACACATATATTCATTCCTCTCTTTGAAGCCACGGTAGATCTCGATTCCTCATCTCAG	1758
QY	973	CTCCATGTTCTTCCGGAAGCAGGTTGTAGGGTTGACCTGTTGATGATGAAGTAAGCA	1032
DB	1759	CTCCATGTTTTTTGAAGCAGGTTGTGGATTGTGATTTGATGATGAAGTAAGCA	1818
QY	1033	GAAGGGCTCAACAAAGACACATGCCACACCTGAACAGTGGACCAATGTGTTCAACCT	1092
DB	1819	GAAGAGCTCCCAACAAACACATGCCACTCCAGTCAATGCACTTAACGCAATCAATCCT	1878
QY	1093	GCATTTTCATATTATGGTACTACTGCTATGCTTAATTTTACCTTAACCAAGCTGGT	1152
DB	1879	GCATTTTCGATATTATGCTACTATTGTTATGCTAACCTCTATGTTTAAATTAAGCTTCA	1938
QY	1153	GAGTCAAGGGAATGACCACATCAAAATTCGTCACATGCTGGAGAGGCTGGAGATGTT	1212
DB	1939	GAGTCAAGGGAATGACATCAATTCGTCACATGCTGGAGAGGCTGGAGATGTT	1998
QY	1213	GATCACTTGGCAGGACATTTCTTCTCTGTGCACAAACATATCACATGGAATTAATCTAAG	1272
DB	1959	GACCACTGGCTGCTACGTTTCTAAACATGCCATAGCATCGCACATGGAATCAATCTGCA	2058
QY	1273	AAGTCTCTGTGCTTCTGATCTGTTACTATCTTGGTTCAGATGTTGCTGGCGATGTC	1332
DB	2059	AAGTCTCTGTGCTTCTGATCTGTTACTTACTCTGCTGCCCAAGATTTGCTGGCCATG	2118
QY	1333	TTGAGCAACACTCTTTTATTTCTTGATCATCATCGAACCTTTTCCAACTCTTCCAA	1392
DB	2119	CTGAGCAACACTCTTTTGTCTTAGATTACCCCGAATCCCGTTTCTGTGTTTCTTCTTA	2178
QY	1393	CGAGGTCTGATGCTCATTTATCTACGGATGACCTTTTGCAAAATTCACCTGACAAAAA	1452
DB	2179	AGAGGTCTCAATGTTTCTGCTTACTGATGACCCCTTTCAGATTCATTAACATAAGAA	2238
QY	1453	CCATTGGTGGAAATACAGCAATTCGCTTTCGCTGGAGAGCTCAGTTCTGTGATTTA	1512
DB	2239	CCTCTCGTGGAGAGTATAGCATAGCTGCATCAGTATGGAAGCTGAGTGGCTGTGACCTG	2298

QY 1513 TCGGAATTCGCGAGGAACCTCTGTTTACCAATCTGGGTTTTTACATGCTCTCAAGGCGAC 1572  
 Db 2299 TCGGAGATAGCTCGTAACCTCAGTGACCACTGAGGTTTCTACACGCCCTGAAGTCGCAC 2358  
 QY 1573 TGGATTGGTAAGAACTACTTCAAAAGAGGACCTGCTGGAAATGATATTCACAGAACCAAT 1632  
 Db 2359 TGGATTGGAAGATTTACTTACAAAGAGGACCTGATGGAAACGACATTCACAAACAAAC 2418  
 QY 1633 GTACCCACATCAGGGTTCAATTTAGAGAGATGATCTGAGAGAAATGAATGAACCTAGTG 1692  
 Db 2419 GTCCACACATAAGGGTGGAGTTCCGTGACAGATCTGGAAGAGGAGATGCAACAGGTT 2478  
 QY 1693 TACTCTGACAATG 1705  
 Db 2479 TATCTGGGCAAGG 2491

## RESULT 6

AAZ23395 ID AAZ23395 standard; DNA; 2880 BP.

XX AAZ23395;

XX 13-DEC-1999 (first entry)

XX A. thaliana AMP-deaminase DNA.

XX AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;  
 KW inhibitor; resistance; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers  
 FT 78..2597  
 CDS /\*tag= a  
 FT /\*product= "AMP-deaminase"

XX WO950400-A1.  
 XX 07-OCT-1999.

XX 25-MAR-1999; 99WO-EP02016.

XX 01-APR-1998; 98DE-1014512.

XX (BAD1 ) BASF AG.

XX Lerchl J, Reindl A;

XX WPI; 1999-580759/49.

XX P-PSDB; AAY33456.

XX DNA encoding AMP deaminase, assay systems for identifying inhibitors  
 and transgenic plants -

XX Claim 1; Fig 2; 45pp; German.

XX This invention describes a novel Arabidopsis thaliana AMP-deaminase  
 (adenosine triphosphate aminohydrolase, EC 3.5.4.6) which has herbicidal  
 activity. The AMP deaminase DNA sequence, can be introduced into pro- or  
 eukaryotic cells, with the relevant control elements to control  
 transcription and translation in the cell. An expression cassette  
 derived from the products of the invention is used to transform plants.  
 CC It is useful for producing a test system to identify inhibitors of AMP  
 CC deaminase. Plants containing the expression cassette are useful for the  
 CC production of the AMP deaminase. The plants have an increased resistance  
 CC against inhibitors of the AMP-deaminase through strengthened expression  
 CC of the AMP deaminase DNA sequence. The expression cassette is useful for  
 CC producing plants with an increased content of AMP. This sequence encodes  
 CC the A. thaliana AMP-deaminase described in the invention.

XX Sequence 2880 BP; 829 A; 600 C; 647 G; 804 T; 0 other;

Query Match 54.5%; Score 1041.8; DB 20; Length 2880;  
 Best Local Similarity 76.0%; Pred No. 2.5e-182;  
 Matches 1286; Conservative 0; Mismatches 407; Indels 0; Gaps 0;

QY 13 GTTGCTCCATGGGAGAGAGGTCTATAAATGACCCCTGACTTCCAAAAACCTAAACCCCAAC 72  
 Db 876 GTTGCAACATGGGAAAAAGAGTCTATCTGATCTCTAGTCTCCAAAGCCTAATACAGAG 935  
 QY 73 CCGTTCACTTATGTCCTGAACCAAGTCAGACATGTTTTCCAAACTGTTGATGCGGTT 132  
 Db 936 CCAATTGCACACTATCTCTCAGGAAAAATCTGATCAATGTTTGTGAGATCAAGATGGGTT 995  
 QY 133 ATCCATGTTTATCGGATATAAAGATTGTACGGAGAGCAATTTATCTCTGCTGGCTGATCTACA 192  
 Db 996 GTCCACGTTGTTGCAATTAAGATGCAAAAGAGATCTCTTCCCGGTAGCTGATGCCACA 1055  
 QY 193 ACCTTCTTCACTGACTTGCATTATATCTCCGAGTAACGGCTGCAGGAAACACAAGAACT 252  
 Db 1056 CGGTTTTCACTGACTTGCATCAGTACTCAAAGTTCATAGCTGCAGGAAACATCCGACT 1115  
 QY 253 GTTCCCAATATCGGTTAAATCTTCTGAGCATAAGTTTAAATTCATCTGATGTTAAAT 312  
 Db 1116 TTGTCCACCGTCGACTAGTCTCTCTAGAACAGAAATTTAATCTCCATTTGATGCTTAAT 1175  
 QY 313 GCGGATAGGGAATTTCTTGGCCAGAGACTGCCCCACATCGTGATTTTACAAATGTCAGG 372  
 Db 1176 GCGGATAAAGAAATTTCTGCTCAAAAAGTGCACACATCGTGATTTTATACGTTAGG 1235  
 QY 373 AAGTTGACACTCATGTTTCATCATTCAGCATGCAATCAAAAAACATCTGTTGAGGTTTC 432  
 Db 1236 AAAGTCGACACTCATGTCATCATTCAGCTTGCATGAACACAGAAACACCTTTTAAGTTT 1295  
 QY 433 ATAAATCCAAACTAAGAAAGAACCTGATGAGGTGGTCATTTTCAGAGATGGTACTTAT 492  
 Db 1296 ATTAAGTCAAAAGCTCCGGAAGAAACCCGATGAGGTTGTAATATTCGAGATGGAACATAT 1355  
 QY 493 ATGACTTTAAAGAGGTTTTTTCAGAGCTTGAGCTTAACTGGGTATGATCTGATGTTGAT 552  
 Db 1356 TTGACCTTGAGAGAGTTTTCAGAGCTGCTGATCTGATGATATGACCTGAACTGCAC 1415  
 QY 553 TTGCTAGATGTCATGACAGAAAAAGCACATTTTCATCGTTTTTGACAAAATTCATATAAA 612  
 Db 1416 CTTTTCGATGTTTCATGACAGAAAAAGTACCTTTTCATCGTTTTTGAAGTTCAACCTAAAG 1475  
 QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAACACAGATATCTTATT 672  
 Db 1476 TATAACCTTTGTGGTCAAAAGTAGGCTTAGGGAGATTTTCTTAAACAGGATAATCTCATC 1535  
 QY 673 CAAGGCCGTTTTTCTGCTGAGTTGACAAAGCAAGTTTTTCTCTGACCTTCTCTAGCAAA 732  
 Db 1536 CAAGTCCGATTTCTTGGTGAGATAACAAGCAAGTCTTCTCTGACCTTGAAGCTAGTAAA 1595  
 QY 733 TATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA 792  
 Db 1596 TATCAGATGGCTGAATACAGATAATCTATATGTCAGAAAAAATGAGCGAGTGGACCAA 1655  
 QY 793 CTTGCAAGTTGGATAGTGAAACAAATGAAATTCAGTGGAAATGTTGTCTGGCTGGTTTCAG 852  
 Db 1656 CTCGCTAGTTGGATTGTGAACAAATGATCTATACAGTGAGAAATGTTGTCTGGTTAAATTCAG 1715  
 QY 853 ATTCACCGTTTATATAATGTGTACAGGAAATGGGTATCGTTTACATCATTTCCAAAATCTT 912  
 Db 1716 CTCACCGCTTGTACACATTTTACAGGACATGGGTATTTGACATCGTTCCAGATATC 1775  
 QY 913 CTTGACACATTTTCGTTCTCTCTTTTTCGAGGTTTACTATTGATCCAGCTTTCACACCCACAG 972  
 Db 1776 CTTGACATATATTTCTCTCTCTGTTTGAAGCCACCGTAGATCTCTGATTTCCCATCTCAG 1835  
 QY 973 CTCCTATGTTCTCCTCAAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAATGAACCA 1032  
 Db 1836 CTCCTATGTTTTTTTGAAGCAGGTTGTTGATTTGATTTGATGATGAAGAAACCACT 1895

```
QY 1033 GAAAGCGTCCAAACAAACACATGCCCCACACCTGACAGTGGACCAATGTGTCAACCT 1092
Db 1896 GAAAGCGTCCCAACAAACACATGCCCCACACCTGACAGTGGACCAATGTGTCAACCT 1955
QY 1093 GCATTTTCATATATATGCGTACTGCTATGCTAACTATTACACCTAAACAAAGCTGGT 1152
Db 1956 GCATTTTCGATATATGCTACTATGTTATGTTAACTTAACTTAACTTAACTTAACTTAA 2015
QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCCTGTCACATGCTGGAGAGCTGGAGATGT 1212
Db 2016 GAGTCAAGGGAATGACTACTATCACTACAGCTTACGACACATCTTGGAGAGCTGGTACAT 2075
QY 1213 GATCACTTGGAGCGACATTTCTTCTGTCACACATATACATGGAATTAATCAAGG 1272
Db 2076 GACCACTTGGCTGCTACGTTTCTAACATGCCATAGCATGCGCATGGAATCAATCTGGA 2135
QY 1273 AAGTCTCTGCTTTCAGTACTTGTACTATCTTGTGATGCTGCGGATGCTGCGGATGCTGCGG 1332
Db 2136 AAGTCTCTGCTTTCAGTACTTGTACTATCTTGTGATGCTGCGGATGCTGCGGATGCTGCGG 2195
QY 1333 TTGAGCAACACTCTTATTTTGAATATCAATGCAACCTTTTCCAAAGTCTTCCAA 1392
Db 2196 CTGAGCAACACTCTTATTTTGAATATCAATGCAACCTTTTCCAAAGTCTTCCAA 2255
QY 1393 CGAGTCTGAATGCTCATTTATCTAGGATGACCTTTGCAAAATTCACCTGACAAAGAA 1452
Db 2256 AGAGTCTCAATGTTTCTGCTACTGATGACCCCTTTCAGATTCACCTTAACTAAAGAA 2315
QY 1453 CCATTTGGTGGAGAAATACAGCAATGCTGCTGCTGGAAGCTCAGTTCTTGTGATTTA 1512
Db 2316 CTTCTGCTGGAAGATATAGCATAGCTGATGATGATGATGATGATGATGATGATGATG 2375
QY 1513 TCGGAAATGCGAGAACTCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGCGCAC 1572
Db 2376 TCGGAGATAGCTCGTAACTCAGTGTACCACTGATGATGATGATGATGATGATGATGATG 2435
QY 1573 TGGATTTGTAAGACTACTTCAAAAGAGAGCTGCTGGAATATGATTTCAAGAACCAAT 1632
Db 2436 TGGATTTGTAAGAGATTTACTACAAAGAGAGAGCTGATGGAACGACATTTCAAAACAAAC 2495
QY 1633 GTACCGCACATCAGAGTTCAATTTTAGAGAGATGATCTGAGAAATGAAATGAACTAGT 1692
Db 2496 GTGCCACACATAGAGGTGGAGTTCCGTGACACGATCTGGAAGAGAGATGACACAGTT 2555
QY 1693 TACTCTGACAATG 1705
Db 2556 TATCTGGCAAGG 2568
```

## RESULT 7

ABZ13910  
ID ABZ13910 standard; DNA; 1803 BP.

XX AC ABZ13910;

XX DT 21-JAN-2003 (first entry)

XX DE Arabidopsis thaliana stress regulated gene SEQ ID NO 1715.

XX DX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

XX KW Arabidopsis thaliana.

XX OS Arabidopsis thaliana.

XX PN WO200216655-A2.

XX PD 28-FEB-2002.

XX PF 24-AUG-2001; 2001WO-US26685.

XX PR 24-AUG-2000; 2000US-227866P.

XX PR 26-JAN-2001; 2001US-264647P.

XX PR 22-JUN-2001; 2001US-300111P.

PA (SCRI ) SCRIPPS RES INST.  
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.

PI Harper JF, Kreps J, Wang X, Zhu T;

XX WPI; 2002-304127/34.

DR Identifying a stress condition to which a plant cell has been exposed  
XX and producing plants with increased tolerance to these abiotic stresses

PS Claim 144; SEQ ID NO 1715; 577bp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant  
CC cell has been exposed, comprising:  
CC (a) contacting nucleic acid representative of expressed polynucleotides  
CC in the plant cell with an array or probes representative of the plant  
CC cell genome; and  
CC (b) detecting a profile of expressed polynucleotides in the plant cell  
CC characteristic of a stress response. The method is useful in the  
CC production of transgenic plants, cells and seeds and in producing plants  
CC with increased tolerance to abiotic stress. The present sequence is that  
CC of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
CC in methods of the invention.

CC Note: The sequence data for this patent is not represented in the printed  
CC specification but is based on sequence information supplied to Derwent by  
CC the European Patent Office.

XX SQ Sequence 1803 BP; 524 A; 384 C; 378 G; 517 T; 0 other;

Query Match 53.8%; Score 1028; DB 24; Length 1803;

Best Local Similarity 76.2%; Pred. No. 8.4e-180;

Matches 1265; Conservative 0; Mismatches 395; Indels 0; Gaps 0;

QY 13 GTTCCTCATGGGAGAGGAGTCTATAATGACCCCTGTACTCCAAACCTTAACCCCAAC 72

Db 100 GTTCACCATGGGAAAAGAGTCTATCTGATCTCTAGTACTCCAAAGCTTAATACAGAG 159

QY 73 CCGTTCATTTATGTCCTGAAACCAAGTCAGAGCATGTTTCCAAACCTGTTGATGGCGTT 132

Db 160 CCATTTGCACACTATCTCTCAGGGAATCTGATCATTTGTTTGTAGATGCAAGATGGGTT 219

QY 133 ATCCATGTTTATGCGGATAAAGATTGTACGAGAGCATTTATCTGTGGTGTGCTTACA 192

Db 220 GTCCACGTTTGTCAAATTAAGATGCAAAAGACATCTCTCCCGGTAGTGTGCCACA 279

QY 193 ACCTTCTCACTGACTTGCATTATTTCTCGAGTAACGGCTGCGAGGACACAGAACT 252

Db 280 GCGTTTTTCACTGACTTGCATCAGTACTCAAAGTCTAGCTGCAGGAAACATCCGGACT 339

QY 253 GTCTGCCATAATCGTTTAAATCTTCTTGAGCATAAATTTAAATCCCATCTCATGTTAAAT 312

Db 340 TTGTGCCACCGTGCAGTAGTGTCTCTAGACAGAAATTTAATCTCATTGATGCTTAAT 399

QY 313 GCGGATAGGAATTTCTTGCCCAAGACACTGCCCCACATCGTGATTTTAAATGTCTCAGG 372

Db 400 GCGGATAAAGAAATTTCTTGCTCAAAAAGTGCAACACATCGTGATTTTATAACGTTAGG 459

QY 373 AAGTTGACACTCATGTTTCATCTTCAAGATGCAATGAATCAAAAACATCTGTTGAGGTTT 432

Db 460 AAGTCGACACTCATGTTGTCATCTTCAAGTGTGATGAACCAAGAAACACCTTTTAAGGTTT 519

QY 433 ATAAATCCAAACTAAGAAAAGAACCTGATGAGTGTGCTATTTTTCAGAGATGGTACTTAT 492

Db 520 ATTAAGTCAAAGCTCCGGAAGAAACCCGATGAGTTGTATATTTCCGAGATGGAACATAT 579

QY 493 ATGACTTTTAAAGAGGTTTTTGTAGAGCTTGGACTTAACTGGGTATGATCTGAAATGTTGAT 552

Db 580 TTGACCTTGAGAGAAGTTTTTGTAGAGCTTGGATCTGACTGGATATGATGACCTGAACGTCGAC 639

QY 553 TTGCTAGATGTCATGTCAGACAAAGACATTTTCATCTGTTTGTGACAAATCAATCTAAAA 612

Db 640 CTTTGTGATGTTTCATGTCAGACAAAGTACCTTTTCATCTGTTTGTGATAAGTTTCAACCTAAAG 699

```
QY 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCTCAAAACAGATAATCTTATT 672
Db      |||
QY 700 TATAACCCCTGTGGTCAAGTAGGCTTAGGAGATTTTCTTTAAACAGGATAATCTCATC 759
Db      |||
QY 673 CAAGGCGGTTTCTTCTGTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTGACAAA 732
Db      |||
QY 760 CAAGGTCGATTTCTTGTGAGATAACAAAGCAAGTCTTCTCTGACCTTTGAAGCTAGTAAA 819
Db      |||
QY 733 TATCAGATGCGCAAAATATAGGATTTCAATCTACGGAAGGAACAGAGTGAATGGGACCAA 792
Db      |||
QY 820 TATCAGATGCTGAAATACAGAAATATCTATATATGGCAGAAATAATGAGCGAGTGGACCAA 879
Db      |||
QY 793 CTTGCAAGTTGGATAGTGAACAAATGAATTCACAGTGGAAATGTTCTGTGGCTGGTTCTAG 852
Db      |||
QY 880 CTCGCTAGTTGGATTTGACAAATGATCTATACAGTGAGATGTTCTGTGGTTAAATCTAG 939
Db      |||
QY 853 ATTCCACGCTTATATATATGTTGTAACAGGAATGGGTATCGTTACATCATTTCCAAAATCTT 912
Db      |||
QY 940 CTCGCCACGCTTGTAACAATTTACAAGGACATGGGTATTTGTGACATCGTTCCAGAAATATC 999
Db      |||
QY 913 CTTGCAACAAATTTCTGTTCTCTCTTTTGTAGGTTACTATTGATCCAGCTTCACACCCACAG 972
Db      |||
QY 1000 CTGGACAATATATTCATTCTCTGTTTGAAGCACCGGTAGATCCTGATTCCTCCATCTCAG 1059
Db      |||
QY 973 CTCATGTCCTCTGAGCAGAGTTGTAGGTTTGGACCTGTTGATGATGAAAGTAAACCA 1032
Db      |||
QY 1060 CTCATGTTTCTTGAAGCAGGTTGTTGGATTTGATTTGTTGATGATGAACCAACCT 1119
Db      |||
QY 1033 GAAAGGCTTCAACAAAGCAGATGCCACACCTGAAACAGTGGACCAATGTTTCAACCT 1092
Db      |||
QY 1120 GAAAGACGTCACCAAAACACATGCCACCTCCAGCTCAATGGACTAAACGATTCATCTCT 1179
Db      |||
QY 1093 GCATTTTCATTTATGCTGTTACTACTGTTACTGTTACTTATTCACCTTCAACAGCTGCGT 1152
Db      |||
QY 1180 GCATTTTCGTTATGTTACTACTGTTACTGTTACTGTTACTGTTACTGTTACTGTTACTG 1239
Db      |||
QY 1153 GAGTCAAAAGGAATGACCACTATCAATTCCTGTCATGCTGAGAGGCTGAGAGATGTT 1212
Db      |||
QY 1240 GAGTCAAAAGGATGACTACTATCACTGACGACCACTTCGAGAGGCTGAGATTT 1299
Db      |||
QY 1213 GATCATTGTCAGGAGCATTTCTTCTGTCATCAACATATCAATGAATTAATCTAAG 1272
Db      |||
QY 1300 GAGCACTTGGCTGTACGTTTCTTAAACATGATCAGATGATGATGATGATGATGATGATG 1359
Db      |||
QY 1273 AGCTCTCTGCTGCTGAGTACTGTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1332
Db      |||
QY 1360 AGTCTCTGCTGCTGAGTACTGTTACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1419
Db      |||
QY 1333 TTGAGCAACAACTCTTATTTCTTGTACTATCATCGCAACCTTTTCCAAAGTTTCTTCCAA 1392
Db      |||
QY 1420 CTGAGCAACAACTCTTGTGTTTCTAGATTACCAACCGGACCGTTTCTGTTGTTTCTT 1479
Db      |||
QY 1393 CGAGGCTGAATGTTCTCATTTATCTAGGATGACCTTTTGAATTCACCTGACAAAGAA 1452
Db      |||
QY 1480 AGAGGCTCAATGTTTCTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1539
Db      |||
QY 1453 CCATTGGTGAAGATACAGATGCTGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1512
Db      |||
QY 1540 CCTCTGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1599
Db      |||
QY 1513 TGGCAATTTGCGAGGAACTCTGTTTACCAATCTGTTGTTTTCATGCTCTCAAGGGGCGAC 1572
Db      |||
QY 1600 TGGGATAGTCTGTAATCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCTAGTCT 1659
Db      |||
QY 1573 TGATTTGGTGAAGATCTCTTCAAAAGAGGACCTGCTGGAATGATATTCACAGAACCAAT 1632
Db      |||
QY 1660 TGGATTTGAAAAGATTTACTACAAAGAGGACCTGATGGAACGACATTCACAAAACAAAC 1719
Db      |||
QY 1633 GTACCGCACATCAGGGTTTCAATTTAGAGAGATGATCTGGA 1672
Db      |||
QY 1720 GTGCCACATAGGGTGGAGTTCCGTGACACGGTATGGA 1759
Db      |||
```

```
RESULT 8
AAF25496
ID AAF25496 standard; DNA; 2482 BP.
XX
AC AAF25496;
XX
DT 15-MAY-2001 (first entry)
XX
DE Nucleotide sequence of a corn adenosine deaminase enzyme.
XX
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
   inosine; ss.
XX
OS Zea mays.
XX
FH Key Location/Qualifiers
FT CDS 596..2482
FT FT /*tag= a
FT FT /product= "adenosine deaminase"
XX
PN WO200109305-A2.
XX
PD 08-FEB-2001.
XX
PR 28-JUL-2000; 2000WO-US21009.
XX
PR 30-JUL-1999; 99US-0146473.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI (PION-) PIONEER HI-BRED INT INC.
XX
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX
DR WPI; 2001-159866/16.
XX
DR P-PSDB; AAB31956.
```

New polynucleotides encoding AMP deaminase or adenosine deaminase used e.g. for designing or identifying herbicides that inhibit the enzyme activities, and as probes for genetic or physical mapping

Claim 2; Page 63; 72pp; English.

The present sequence encodes an andenosine deaminase. The specification also describes adenosine deaminase. These enzymes convert adenosine to inosine. Mutations in these genes cause disruptions in then salvage and catabolism of adenosine and AMP. In humans, this may lead to death of white blood cells, which causes severe immunodeficiencies. The AMP deaminase and adenosine deaminase may be used to prepare antibodies to these proteins, and to design or identify herbicides that inhibit their enzyme activities. The polynucleotides are used as probes for genetically and physically mapping genes that they compose, and as markers for traits linked to those genes, where such information may be used in plant breeding to develop lines with desired phenotypes. The nucleic acid fragments may be used to isolate cDNAs and genes encoding homologous proteins from the same or other plant species, and in polymerase chain reaction (PCR) protocols to amplify longer nucleic acid fragments encoding homologous genes from DNA or RNA. These may also be used to create transgenic plants in which the polypeptides are overexpressed or suppressed, and as probes in direct fluorescent in situ hybridisation (FISH).

Sequence 2482 BP; 688 A; 539 C; 569 G; 686 T; 0 other;

Query Match 50.4%; Score 962.4; DB 22; Length 2482;  
Best Local Similarity 78.7%; Pred. No. 9.5e-168;  
Matches 1149; Conservative 0; Mismatches 311; Indels 0; Gaps 0;

```
QY 13 GTTGCTCCATGGAGAGGAGGTCATAAATGACCCCTGACTCCAAAACCTTAACCCCAAC 72
Db 1023 GTTGCTCCGTTGGATTAAGAAGTTATATCCGACCCAGCAGACCCAGCTTAACCCAGAT 1082
QY 73 CCGTTCACTTATGTGCTGCAACCAAGTCAGAGCATGTTTTTCCAAACTGTTGATGCGGTT 132
```

Db 1083 CCAATTTTATACATCTCTGAGGAAATTCGTATCATATATTTGAAATGCAAGATGGGTT 1142  
QY 133 ATCCATGTTTATCGGATAAAGATTGTACGGAGAGCAATTTATCTCTGGCTGGCTGATGCTACA 192  
Db 1143 ATTCTGTTATATCCAGATAGAGATGCAAAAGAGAGCTTTTCTCTAGCCGATGCAACT 1202  
QY 193 ACCTTCTTCACTGACCTTGATATATTTCTCCGAGTAAACGCTGCGAGGAACAACAAGAACT 252  
Db 1203 ACATTTTTCACGTATCTTCACTTACTTCGAGTCATAGCAGCAGGGAATATAAGAACT 1262  
QY 253 GTCTGCATAATCGGTTTAAATCTTCTTGAGCATAAAGTTTAAATTCCTCATCTGATGTTAAAT 312  
Db 1263 TTATGCCATCATAGGCTCAATCTTCTAGAACAAAATTCATCTTCAATTTGATGCTAAAT 1322  
QY 313 GCGGATPAGGAATTTCTTCCAGAGAGACTGCCACATCGTGATTTTCAATGTCCAGG 372  
Db 1323 GCGGATPAGGAATTTCTTCTCAGAGAGTGCTCCACATCGAGACTTCTATAATGTAGA 1382  
QY 373 AGGTTGACACTCATGTTCAATCATTTGAGCATGATGAATCAAAAACATCTGTTGAGGTTTC 432  
Db 1383 AAGTTGATACTCATGTCACCACTCAGCATGATGAATCAGAAACATCTTTTAAAGTTTC 1442  
QY 433 ATAAATCCAACTAAGAAAGAACCTGATGAGTGCTCATTTTCAGAGATGCTATAT 492  
Db 1443 ATAAAGTCAAGCTGAGAAAGAGCCTGATGAGTTGTAATATTTTCGAGATGGGACATAT 1502  
QY 493 ATGACTTTTAAAGAGGTTTTTTCAGAGCTTTGAGCTTAACTGCGGTGATGATCTGAATGTGAT 552  
Db 1503 CTAACGTTGGAAGAGGTTTTCAAGAGTTTAGATTTTGTCTGGATATGACCTCAATGTGAC 1562  
QY 553 TTGCTAGATGTCATGACAGACAAAAGACATTTTCAATCGTTTTTGAACAAATCAATCTAAA 612  
Db 1563 CTTTGGAGCTTCACGACACAGAGTACTTTTTCATCGCTTTGATAAGTTCAATCTTAA 1622  
QY 613 TACAATCCATGTCGCAAGTAGGCTCAGAGAAATTTTCTCAACAAAGATATCTTAT 672  
Db 1623 TACAATCTTGGCTCAAGTAGGCTCAGGAGATATTTCTTAAGCAGGATATCTCAT 1682  
QY 673 CAAGGCGTTTTCTGCTGAGTTGCAAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAA 732  
Db 1683 CAAGGTCGTTTTCTGGTGAGTTAACTAAGCAAGTTGTTTCAGATCTCTGCTCCAGTAA 1742  
QY 733 TATCAGATGGCAGATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAA 792  
Db 1743 TATCAGATGGCTGAATATAGAAATATCAATATATGTTAGGAAGCAAAAGTGAAGTGGACCAA 1802  
QY 793 CTTGCAAGTTGGATAGTGAACATGAAATTCACAGTGGAAATGTTGCTGGCTGGTTGAC 852  
Db 1803 CTAGCCAGTTGGATAGTGAATATGATTTGTACGCGGAATGTCGTTTGGTTGATTCAG 1862  
QY 853 ATTCCACGCTTATATAATGTACAGGAATGGGTATCGTTACATCATTTCCAAATCTT 912  
Db 1863 CTTCCACGCTTGTACATGTGTACAGAAATGGGAATTTGTACATCATTTCCAGAACATG 1922  
QY 913 CTGCAACAACTTTTCTGCTCTTTTGGAGTTACTTATGATCCAGCTTTCACACCCACAG 972  
Db 1923 CTCGCAATATTTTCAATCCACTTTTGGAGTCACTGTCAACCCAGATTCACATCCTCAG 1982  
QY 973 CTCATGCTTCTGAGCAGGTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACCA 1032  
Db 1983 CTGCAATGTTTCTGAAACAGGTTGTGGGTTGGATTTGGTTGATGATGAAGCAACCT 2042  
QY 1033 GAAAGCGCTCCAAAGACATGCCACACCTGAAACAGTGGACCAATGTTTCAACCT 1092  
Db 2043 GAAAGCGCCAAACAAACATGCTACACCTGAGCAATGACTATGTTTTCATCG 2102  
QY 1093 GCATTTTCAATATGCGTACTGCTATGCTAACTTATCCACCTTAAACAGCTGGT 1152  
Db 2103 GCATTTTCAATATGCTTACTTATGCTTATGCAAACTTTTACACCTTAAACAGCTTGA 2162  
QY 1153 GAGTCAAGGGAATGACCACTATCAATTTCCGTCACATGCTGGAGAGCTGGAATGTT 1212

Db 2163 GAATCAAGGGAATGACAAACATCAAAATTCGTCCACATTTCTGGAGAGCTGGTGATATT 2222  
QY 1213 GATCACTTGGCAGGACATTTTCTCTCTGTCTCAACATATCACATGGAATTAATCTAAGG 1272  
Db 2223 GACCACCTTGGCAGCAACCTTTCTCAGGCTCACAACATTTGCACATGGAATCAATTTGAAA 2282  
QY 1273 AAGTCTCTGCTGCTCAGTACTTGTACTATCTTGTGATGATGCTGCGGATGTCCTCA 1332  
Db 2283 AAATCTCTGCTGCTCAATATTTATTTAGCCAGATTTGGGCTGGCAATGTCTCT 2342  
QY 1333 TTGACCAACAACTCTCTTATTTCTGACTATCATCGCAACCTTTTCCAAACGTTCTTCCAA 1392  
Db 2343 TTGACATAACTCCCTTATTTCTGACTATCATCGGAATCTTTTCCAAATGTTCTCTTA 2402  
QY 1393 CGAGTCTGAATGTCTATTTCTACGGATGACCTTTTGCAAATTCACCTGACAAAGAA 1452  
Db 2403 CGGGTCTGAATGTGTCACTTTCTACTGATCTCTCCTCAAAATTCCTTAAACAAAGAA 2462  
QY 1453 CCATTTGGTGGAGATACAG 1472  
Db 2463 CCATTTGGTGGAGATATAG 2482

RESULT 9  
AAF25489  
ID AAF25489 standard; DNA; 1816 BP.  
XX  
AC AAF25489;  
XX  
DT 15-MAY-2001 (first entry)  
XX  
DE Nucleotide sequence of a rice AMP deaminase enzyme.  
XX  
KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
KW inosine; ss.  
XX  
OS Oryza sativa.  
XX  
FH Key Location/Qualifiers  
FT CDS 3..1481  
FT /\*tag= a  
FT /product= "AMP deaminase"  
XX  
PN WO200109305-A2.  
XX  
PD 08-FEB-2001.  
XX  
PF 28-JUL-2000; 2000WO-US21009.  
XX  
PR 30-JUL-1999; 99US-0146473.  
XX  
PA (DUPO) DU PONT DE NEMOURS & CO E I.  
XX (PION-) PIONEER HI-BRED INT INC.  
PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;  
XX  
XX WPI; 2001-159866/16.  
DR P-PSDB; AAB31949.  
XX  
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used  
PT e.g. for designing or identifying herbicides that inhibit the enzyme  
PT activities, and as probes for genetic or physical mapping  
XX  
PS Claim 2; Page 44-45; 72pp; English.  
XX  
CC The present sequence encodes an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for

CC Genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

XX  
SQ Sequence 1816 BP; 566 A; 344 C; 381 G; 525 T; 0 other;

Query Match 50.2%; Score 959; DB 22; Length 1816;

Best Local Similarity 79.4%; Pred. No. 3 9e-167;

Matches 1136; Conservative 0; Mismatches 295; Indels 0; Gaps 0;

QY 261 TAAATCGGTTAAATCTCTTTGAGCATAAAGTTTAAATCCATCGATGTTAAATGCGGATAG 320

DB 8 TAAACGTTTAAATCTCTAGACAGAAATTCATCTTCAATTTGATGGTCAATGCCGATAG 67

QY 321 GGAATTTCTGCCAGAGAGCTGCCCATCATCTGATTTTACAAATGTCAGGAGTTTGA 380

DB 68 AGAACTACTTGTCTCGAAAGAGCTGCCCATCGGAGCTTCTACAATGTCAGGAGGTTGA 127

QY 381 CACTCATGTTTCATCATCTCAGCATGCAATGAATCAAAACATCTGTTGAGGTTTCATAAAATC 440

DB 128 TACTCATGTTTCATCATCTCAGCATGCAATGAATCAAAACATCTGTTGAGGTTTCATAAAATC 187

QY 441 CAAACTAAGAAAGAACCTGATGAGTGGTTCATTTTTCAGAGATGGTACTTATATAGCTTT 500

DB 188 CAAAGTTGAGAAAGAACCTGACGAGGTTGTGATTTTATAGAGATGGTACTTATATAGCTTT 247

QY 501 AAGAGAGTTTTCAGAGCTTGACCTTAATCTGATGATGATCTGAAATGTTGATTTGCTAGA 560

DB 248 TAAGSAGGTTTTCAGAGCTTGACCTTAATCTGATGATGATCTGAAATGTTGATTTGCTAGA 307

QY 561 TGTCCATGACAGAAAGACATTTTCATCGTTTTCAGAAATTCATCTTAAATFACAAATCC 620

DB 308 TGTGATGCGATTAAGATACATCTCATCGTTTTCAGAAATTCATCTTAAATFACAAATCC 367

QY 621 ATGTGCCAAAGTAGGCTCAGAGAAATTTTCTCAAAAGATAATCTTATTCAGAGCCG 680

DB 368 TTGTGCCAATCCCGCTGAGGAGATCTTCTTAAACAGGACAACTTATTCAGAGCCG 427

QY 681 TTTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAATATCAGAT 740

DB 428 ATTCTTGTGTAATGACAAAGAGATTTTCTGATCTTGAAGCAAGTAAATATCAGAT 487

QY 741 GGCAGATATAGGATTTCAATCTACGAGGAAACAGAGTGAATGGACCACTTCAAG 800

DB 488 GGTGATGATAGATATCTATCTATGGAGAAAGAAAGTGAATGGATCAGATGCGAAG 547

QY 801 TTGGATAGTGAACAATGAATTCACAGTGAATGTTGCTGCTGCTGCTGCTGCTGCTGCTG 860

DB 548 CTGGATAGTGAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607

QY 861 CTTATATATGCTACAGGAATGGTATCGTTACATCATCTTCAAAATCTTTCACAA 920

DB 608 GATATACATGATATACAGGAGATGGGAACAATCAATTTCTTCCAGAACTCTTCACAA 667

QY 921 CATTTTCTGCTCTTTTTCAGGTTTACTATTGATCCAGCTTTCACACCAAGCTCCATGT 980

DB 668 TATTTTCTGCTCTTTTTCAGGTTTACTATTGATCCAGCTTTCACACCAAGCTCCATGT 727

QY 981 CTTCTGTAAGAGCTTGTAGGTTTGGACCTGTTGATGATGATGATGATGATGATGATGATG 1040

DB 728 TTTCTTGCACAGCTGCTGGCTGGATTTAGTGGATGATGATGATGATGATGATGATGATG 787

QY 1041 TCCACAGACACATGCCCACTGACAGTGGACCAATGTTTCAACCTGCATTTTC 1100

DB 788 CCCAACAAACACATGCCCTACCTGAGCAATGGACTAATGTTTTCATTCAGCATATGC 847

QY 1101 ATATTATCGTACTACTGCTATGCTAACTATTTCACCCCTAAACAGCTGCGTGAAGTCAA 1160

DB 848 ATATTATGCTACTATTGCTTAACTGCTAACTTGTACACGCTGAACAGCTTCTGCTGAAGTCAA 907

QY 1161 GGGATAGACCACTATCAAAATTCGCTCCACATGCTGGAGAGGCTGGAGATGTTGATCACTT 1220

DB 908 GGGATAGCAACAATCAAACTTCGCTCCACATGCTGGGAGGCTGGAGATATTTGATCATCT 967

QY 1221 GCGAGCGACATTTCTTCTCTGTCACAAACATATACATGGAATTAATCTAAAGAACTCTCC 1280

DB 968 TGCTGACGATTTCTTACTTCTCATATAATTTGCTCAGGGGTTAATTTAAAGAACTCTCC 1027

QY 1281 TGTGCTTCAGTACTTGTACTATCTTGTGTCAGATGCTGCGATGTCCTCCCATTTGAGCAA 1340

DB 1028 TGTCTCTCCAGTATCTGTATTACCTAGCTCAGATGCTTGTGCAATGCTCTCTTTGAGCAA 1087

QY 1341 CAACTCTCTTATTTCTTGACTATCATCGCAACCTTTTTCACAGCTTCTTCCAAAGAGGCTT 1400

DB 1088 CAACTCAATGTTTATTTGATTTATCACGAAACCTTTTCCCAACATTTTCTTAAGAGGCTT 1147

QY 1401 GAATGCTCTCATTTCTACGATGACCTTTTGCAAATTCACCTGACAAAGAACCACTTGT 1460

DB 1148 TAACTGTTCTCTATCAACGATGACCTTTTGCAAATTCACCTGACAAAGAACCACTTGT 1207

QY 1461 GGAAGATACAGCATTTGCTGCTTCTGCTGGAAGCTCAGTTCTTGTGATTTATGCGAAAT 1520

DB 1208 TGAAGATATAGCATCGCTGCTTCTGCTGGAAGCTAAGTTTCATCGACCTATGTAAT 1267

QY 1521 TCGGAGGAATCTGTTTACCAATCTGGGTTTTCACATGCTCTCAAGGCGCACTGATTTGG 1580

DB 1268 TGCTAGGAATTTCTGTTACCACTGCTGTTTCTCTCATAGGCTCAAGTCACTGGAATGG 1327

QY 1581 TAAGAACTACTTCAAAAGAGGACCTGCTGGAATGATATTTCACAAACCAATGTACCGCA 1640

DB 1328 GAGAATCTACTACAAAGAGGTCATGATGGCAATGACATTCACAGACAAATGTTCTCTCA 1387

QY 1641 CATCAGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAAATGAAATAGT 1691

DB 1388 CATCAGGATGAAATTCGACACACTATTTTGGAAAGAAATGGAGCTAAT 1438

RESULT 10

AAL45031

ID AAL45031 standard; cDNA; 914 BP.

XX AAL45031;

XX AC AAL45031;

XX DT 16-MAY-2002 (first entry)

XX A thaliana AMP deaminase coding sequence fragment EST.

XX AMP deaminase; adenosine monophosphate; transgenic plant; EST;

XX herbicide resistance; herbicide; inhibitor; expressed sequence tag; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers

XX CDS 2..880

XX /\*tag= a

XX /product= "AMP deaminase fragment"

XX WO200206319-A2.

XX 24-JAN-2002.

XX 06-JUL-2001; 2001WO-EP07767.

XX 17-JUL-2000; 2000DE-1035084.

XX (AVET ) AVENTIS CROPS SCIENCE GMBH.

XX Schulz A, Streber W, Hanke C, Schmidt F, Schubel A;

XX







PR 07-AUG-2000; 2000US-223323P.  
 PR 05-JUN-2001; 2001US-0873319.  
 XX (GENE-) GENE LOGIC INC.  
 PA (NISB) JAPAN TOBACCO INC.  
 XX Munger WE, Kulkarni P, Getzenberg RH, Waga I, Yamamoto J;  
 XX WPI; 2002-257476/30.  
 XX  
 XX Identifying drugs for and diagnosing benign prostatic hyperplasia, by  
 PT detecting expression levels of one or more genes in prostate cells from  
 PT patient that are differentially regulated compared to normal prostate  
 PT cells -  
 XX  
 XX Disclosure; Page 262-263; 444pp; English.  
 XX  
 CC The invention relates to a method of diagnosing (I) the onset or  
 CC progression of benign prostatic hyperplasia (BPH), or screening (II) for  
 CC or identifying an agent that modulates the onset or progression of BPH.  
 CC The method is based on changes in gene expression in BPH tissue isolated  
 CC from patients exhibiting different clinical states of prostate  
 CC hyperplasia as compared to normal prostate tissue. (I) comprises  
 CC detecting the expression levels of one or more genes in prostate cells  
 CC from the subject that are differentially regulated compared to normal  
 CC prostate cells. (II) comprises preparing a first gene expression profile  
 CC of BPH cells or BPH-like cell population, exposing the cells to the  
 CC agent, preparing a second gene expression profile of the agent exposed  
 CC cells, and comparing the first and second gene expression profiles.  
 CC (I) is useful for diagnosing the onset or progression of BPH. (II) is  
 CC useful for identifying an agent that modulates the onset or progression  
 CC of BPH. The methods are useful to present information identifying  
 CC the expression level in a tissue or cells, by comparing the expression  
 CC level of genes given in the specification in the tissue or cells to the  
 CC level of expression of gene in the database, and displaying the  
 CC expression levels of at least one gene in the tissue or cell sample  
 CC compared to the expression level in BPH. Agents using (II) are useful for  
 CC treating BPH or prostate cancer. AK64106-AK64860 represent human  
 CC benign prostatic hyperplasia gene sequences of the invention.  
 XX  
 XX Sequence 3386 BP; 658 A; 1046 C; 973 G; 709 T; 0 other;  
 SQ  
 Query Match 24.7%; Score 471.6; DB 24; Length 3386;  
 Best Local Similarity 58.3%; Pred. No. 1.3e-77;  
 Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;  
 QY 238 GGGACACAGAAGCTCTGCGCATATCGTGAATCTTCTTGAGCAATGTTAAATTC 297  
 DB 956 GGCCCATAAAGTCATCTGCTACCGCCGGCTGCAGTACCTGAGCTCCAAGTTCAGATG 1015  
 QY 298 CATCTGATGTTAAATGCGGATAGGGAATTTCTTGGCCAGAAGACTGCCCCACATGCTGAT 357  
 DB 1016 CATGTGCTACTCAATGAGATGAAGAGCTGGCGGCCAGAGAAGTGCACACCGAGAT 1075  
 QY 358 TTTTACAATGTCAGNAGTTGACACTCATGTTTCATCTTACGATGATGATGATCAAAA 417  
 DB 1076 TTCTACAACATCGCAAGTGTGACACCCACATCCATCGCTCTGCTGATGAACACAGAAG 1135  
 QY 418 CATCTGTTGAGTTTCATATAATCCAACTAAGAAAGAACCTGATGAGTGTCTATTTTC 477  
 DB 1136 CATCTGCTCGGTTTCATCAAGCGGGCAATGAACGGGCACCTGGAGGAGATCGTGACGTG 1195  
 QY 478 AGAGATGGTACTATATGACTTTAAAGAGAGTTTTTGAGAGCTTGGACCTTAACCTGGGTAT 537  
 DB 1196 GAGCAGGGCCGTTGAACAGACGCTGCGGAGGTTCTTTGAGAGCATGATCTCACGGCTTAC 1255  
 QY 538 GATCTGAATGTTGATTTGCTAGATGTCATGTCAGACAAAAGCACATTTTCATGTTTTCAC 597  
 DB 1256 GACCTGAGTGTGACACGCTGATGTCATGCGGACAGGAACACATTTTCCATCGCTTTCAC 1315  
 QY 598 AATTAATCAATCAATACATCATGTCGCCAAAGTAGCTCAGAGAAATTTTCTCAA 657  
 DB 1316 AAGTTTAATGCCAAATACAAACCTTATTTGGGAGTCCGTCCTCCGAGAGATCTTCAATCAAG 1375

QY 658 CAAGATAATCTTATTTCAAGGCGGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTTCTCTGAC 717  
 DB 1376 ACGGACAAACAGGGTATCTGGGAAGTACTTTGCTCACATCATCAAGGAGGTGATGTGCAGAC 1435  
 QY 718 CTTTCTGCTAGCAAAATATCAGATGCGCAATATAGGATTTCAATCTACGGAGGAACAG 777  
 DB 1436 CTGGAGGAGCAAAATACCAAGATGCAGAGCTGGGCTCTCATTTTCGGGGCTCGAGG 1495  
 QY 778 AGTGAATGGGACCAACTTTGCAAGTTGGATAGTGAACAAATGAATTTGCACAGTGGAAATGTT 837  
 DB 1496 GATGAGTGGGACAAAGCTGGCGCTGGGCGCTCATGCAACCGCTGCACTCCCCAACGCTG 1555  
 QY 838 GTCTGGCTGGTTTCAAGATTTCCACGCTTATATTAATGTGTACAAGGAATGGGTATCGTTACA 897  
 DB 1556 CGCTGGCTGGTGCAGGTGCCCGGCTCTTTGATGTGTACCGGTACCAAGGGCGCAGCTGGCC 1615  
 QY 898 TCATTTCAAAATCTCTTGACAAATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 957  
 DB 1616 AACTTCCAGGAGATGCTGGAGAACATCTCTGCGCACATCTCTGCGAGGCCACTGTGCACCT 1675  
 QY 958 GCTTCCACACCCACAGCTCCATGTCTCTCTGAAGCAGGTTGTAGGTTGGACCTGTGTGAT 1017  
 DB 1676 GCCAGCCACCCGGAAGTCACTCTCTCTTAGACACGTGGATGTTTGTGACAGCGTGGAT 1735  
 QY 1018 GATGAAGTAACACAGAAAGCGCTCCAAACAAGC---ACATGCCCAACCTGAAACAGTGG 1074  
 DB 1736 GATGAGTCCAAAGCTGAAACCATGTCTTCAACCTGGAGAGCCCCCTGCTGAGGCGTGG 1795  
 QY 1075 ACCATGTTTCAACCTGCAATTTTCATTTATGGTACTACTGCTATCTATTAATTAATTC 1134  
 DB 1796 GTGAGGAGGAGCAACCCACCTATGCTTACTACCTGTACTACCTTTTGGCAACATGGCC 1855  
 QY 1135 ACCCTAAACAAAGCTGCGTGAAGTCAAGGAATGACCACTATCAAAATTCGCTCCACATGCT 1194  
 DB 1856 ATGTTGAACCACTTGGCAGCAGCAGAGGGGCTTCCACAGTTTGTGCTGAGGCCACACTGT 1915  
 QY 1195 GGAGAGGCTGAGATGTTGATCATTGTCAGCAGCATTTTCTTCTGTGTCACAAATATCA 1254  
 DB 1916 GGGAGGCTGGGCCCATCCACCATCTGGTGTGAGCTTTCATGCTGGCTGAGAACATTTCC 1975  
 QY 1255 CATGGAATTAATCAAGGAAGTCTCTGCTTCTGATCTTCTGATCTTCTGATCTTCTGATCTT 1314  
 DB 1976 CACGGCTCTCTTCTGCGCAAGGCCCGCTCTGCTGAGTACCTGTACTCTGCTGCGCCAGATC 2035  
 QY 1315 GGTCTGGCGATGTCCCAATTCAGCAACAACTCTTATTTCTTGTACTATCATCGCAACCT 1374  
 DB 2036 GGCAATCGCAATGTCTCGCTCAGCAACAAAGAGCTTCTCTCAGCTATCAACGGAATCG 2095  
 QY 1375 TTTCCAAAGCTTCTTCCAAACGAGGTCTGAATGTCTCATTTATCTACGGATGACCTTTTGCAA 1434  
 DB 2096 CTACCGGAGTACTGTCTCCCGGCTCATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2155  
 QY 1435 ATTCACTGACAAAGAACCAATGTTGGGAAGAAATACAGCATTTGCTTCTGCTGTGGAG 1494  
 DB 2156 TTCCACTTCAACAAAGGAGCGCTGATGAGGAGTACAGCATCGCCACCCAGGTGTGAG 2215  
 QY 1495 CTGAGTTCTTGTGATTTATGGAATTCGAGGAATCTGTTTATCAATCTGGGTTTCA 1554  
 DB 2216 CTGAGCTCTCGATATGTGTGAGTGGCCCGCAACAGCGTGTCTCATGAGCGGCTTCTCG 2275  
 QY 1555 CATGCTCTCAAGGGCCTGGAATTTGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAAT 1614  
 DB 2276 CACAAGGTAAAGAGCCACTGCTGGGACCCAACTATACCAAGGAAGGCCCTGAGGGGAAT 2335  
 QY 1615 GATATTACAGAACCAATGTACCGCATCATAGGGTTCATTTAGAGAGATGATCTGGAGA 1674  
 DB 2336 GACATCCCGGAGCAATGTGCCAGACATCCGCGTGGGCTACCGCTACGAGACCCCTGTGC 2395  
 QY 1675 AATGAATGAAGT 1688  
 DB 2396 CAGGAGCTGGCGCT 2409

## RESULT 13

ABLI7219  
ID ABLI7219 standard; DNA; 2250 BP.  
XX AC ABLI7219;  
XX 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 3130.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical; gene; ds.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX 11-JUL-2000; 2000US-0614150.  
XX (PEXE ) PE CORP NY.  
XX  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signalling and cell-cell  
interactions -  
PT  
PT  
PT  
PS Claim 1; SEQ ID NO 3130; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA  
sequences (ABLI01840-ABLI16175) and the encoded proteins  
(ABBS7737-ABBS72072).  
XX  
XX The sequence data for this patent did not form part of the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 2250 BP; 568 A; 622 C; 545 G; 515 T; 0 other;  
  
Query Match 24.1%; Score 459.4; DB 23; Length 2250;  
Best Local Similarity 57.1%; Pred. No. 2.2e-75;  
Matches 857; Conservative 0; Mismatches 641; Indels 3; Gaps 1;  
  
QY 152 AGATTGTACGGAGACATTATTCCTGTGCTGATGCTACAACTTCTTCACTGACTGC 211  
DB 490 ACAGAACGATGATGATCAATATCAATATCCAGATATGAGCCAAATTCGTAACGACATGC 549  
QY 212 ATTATATTCTCCGAGTAACCGCTGCGAGGAACAACAAGAACTGTCTGCCATAATCGGTTAA 271  
DB 550 AGGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609  
QY 272 ATCTCTTGTAGCATAAGTTTAAATTCATCTGATGTTAAATGCGGATAGGAAATTTCTTG 331  
DB 610 GTTATCTGTCTCCAAAGTATCAGATGCACGTGCTCTCTCAACGAGCTGCGTGTGCTG 669  
QY 332 CCAGAGAGCTGCCACATCGTATTTTACATGTCAGAGAGTTGACATCATGTTTC 391  
DB 670 CCAGAGAGCGGTGCGCATCGGATTTTACAAACACCCGCAAGGTTGACCCACATCC 729  
QY 392 ATCATTGACATGATGAATCAAAACATCTGTTGAGGTTTCATAAAATCCAAACTAAGAA 451

DB 730 ATGCGCATCGTGATGAACCAAAAGCATCTGCTCGCTTCATTTAAAAAGACGCTGAAGA 789  
QY 452 AAGAACCTGATGAGTGGTGCATTTTCAGAGATGGTACTTATATGACTTTTAAAGAGGTTT 511  
DB 790 ACAAAGCAACGAGTGGTCCCGTCACCAATGGCCAGCAATGACATGGCCCAAGGTGT 849  
QY 512 TTGAGAGCTTGGACTTAACCTGGGTATGATCTGAATGTTGATTTGCTAGATGTCCATGCG 571  
DB 850 TCCAGTCATGAACCTGACCACTACGACCTGACCGTCGACATGTTGACGCTCCATGCG 909  
QY 572 ACAAAGCACATTTTCATCGTTTGAACAATCAATCTAAATACAAATCAATGTTGCGCAAA 631  
DB 910 ATCGCAACACGTTCCATCGCTTCGACAAAGTTTAACTCCAAGTACAAATCCCATCGGGAGT 969  
QY 632 GTAGCTCAGAGAAATTTTCTCAACAAGATATCTTATTCAGGCGGTTTCTTGTCTG 691  
DB 970 CGCGCTCAGGAAGTCTTCTCTGAAGACGGAACAATCTCTGAATGGCAAAATCTTTTGAC 1029  
QY 692 AGTTGACAAAGCAAGTTTCTCTGACCTTTCTGTAGCAAAATATCAGATGGCAGAAATATA 751  
DB 1030 AGATCATTAAGGAAGTGCCTTCGATCTGGAGGAGTCCAAGTACCAAGACGCGAGCTGC 1089  
QY 752 GGATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTGCAGATTTGGAATGTA 811  
DB 1090 GTCTCTCCATCTATGGCAAGTCCCGGATGAGTGTACAACTGGCCAAAGTGGGCCATCG 1149  
QY 812 ACATGAATTCACAGTGGAAATGTTGCTGGCTGGTTTCAGATTCACGCTTATATAATG 871  
DB 1150 ACATGAATGTTACAGCTCGAACAATTCGCTGGCTGATCCAGATACCCCGTCTCTTCGACA 1209  
QY 872 TGTAAGGAATGGGTATCGTTACATCATTTCCAAATCTCTTTCGACAAACATTTTCGTTTC 931  
DB 1210 TCTTCAAGTCGACAAAGATGATGAGTCAATTCAGAGATCTTGAACCAATATTTCTCTGC 1269  
QY 932 CTCTTTTGAAGTTACTATTGATCCAGCTTCACACCCAGCTCCATGTTCTTCTGTAAGC 991  
DB 1270 CGCTCTCGAGGCGACAGCCGACCAAGTACGAGTCCGAGTTGATCGCTTCTGCACT 1329  
QY 992 AGTTGTAGGTTGACCTGTTGATGATGAAGTAAACCAAGAGGCTCCCAACAAAGC 1051  
DB 1330 AGCTGATCGGATTCGATTCGTTGGGACGATGATCCAAAGCGGAGAAATCGCTCTTCGACA 1389  
QY 1052 ACA---TGCCCCACACTGAACAGTGGACCAATGTTTCAACCTGCTCATTTTCATATATG 1108  
DB 1390 AGATGTGCCAGCTCCGAGGAGTGGACCTACGAGAGAAATCCGCCCTATGCTACTATA 1449  
QY 1109 CGTACTACTGTATGCTAACTTATTCACCTTAAACAAGTGGTGGTCAAGGGAATGA 1168  
DB 1450 TCTACTATATGTACGCCAACATGACGGTGTGTAACAAAGTTTACAAATCTCGCAATATGA 1509  
QY 1169 CCACATCAATTCGTTCCATGCTGGAGAGGCTGGAGATGTTGATCACTTGGCAGCGA 1228  
DB 1510 ACACCTTTTGTCTGGCGGCCCATTCGCGGAGGCTGGACCGGTCGAGCATCTGGTTGGG 1569  
QY 1229 CATTTCTTCTCTGTCAACAATATCAGATGGAATTAATCTAAGGAAGTCTCTGTGCTTC 1288  
DB 1570 GCTTCTCATGSCCGAGAACATCTCCACGGCTCTCTGTGGCAAGTGGCCGCTGCTG 1629  
QY 1289 AGTACTGTGATCTTGTGATGTTGCTGGGATGTTCCCATTTGAGCAACAACATCTCT 1348  
DB 1630 AGTATCTGTATTATCTGACCCAGATTGGCATCGCCATCTGCGGCTGTGCAACAACATCCC 1689  
QY 1349 TATTTCTTGACATCATCGCAACCTTTTCCAAAGCTTCTTCCAAAGAGTCTGAATGCTCT 1408  
DB 1690 TGTCTCTCAACTATCATCGCAATCCGCTGCCAGAGTATTTCGCCCGTGGCCCTCATCTTT 1749  
QY 1409 CATTTATCTACGATGACCTTTTGCAAAATTCACCTGACAAAGAAACCAATTTGGTGAAGAAT 1468  
DB 1750 CCTCTCCAGGATGATCCCTTGGCAATTCATCCCAAGGAAACCCCTCATGGAGAGT 1809  
QY 1469 ACAGCATGCTGCTCTGCTGTGGAAGCTCAGTCTTGTGATTTATGCGAAATTTGCGAGA 1528  
DB 1810 ACAGCATTTGCGGACAGGTTGGAAGCTCAGCTCATGCGACATGTCGAGCTGGCGCAGGA 1869



853 ATTCACGCTTATATATGTTACAGGAATGGTATGTTACATCATTTCCAAATCTT 912  
Db |||||  
4573 GTTCCAGGANTCATGATGTGTTCCGTTCCAGAAATTTCTTCCACATTTTGGAAATG 4632  
QY  
913 CTTGACAACATTTCTCTCTCTCTTTTGGAGTTACTATTGATCCAGCTTCACACCCACAG 972  
Db |||||  
4633 CTGGAATATTTTCATGCCAGTGTGTTGAGGCCACCATCAACCCAGGCTGACCCAGAA 4692  
QY  
973 CTTCCATGTTCTCTGTAAGCAGGTGTGAGGTTGGACCTGTTGATGATGAAGTAAACCA 1032  
Db |||||  
4693 CTCAGTGTCTCTCAAGCATATCACTGGCTTTGACAGTGTGGATGATGATGATGATGAT 4752  
QY  
1033 GAAAGCC---GTCCACAAGACATGCTCCACACCTGACAGTGGACCAATGTTTCAAC 1089  
Db |||||  
4753 AGTGGCCACATGTTCTCTCCAGAGTCCCAAGCCCCAGGAGTGGACATGGAAAGAT 4812  
QY  
1090 CTTGCACTTTTCTATTTATGCTACTACTGCTATGCTTAATTTTACCCCTAAACAGCTG 1149  
Db |||||  
4813 CCATCTTACACTTACTATGCTACTATGATGATGATGATGATGATGATGATGATGATGAT 4872  
QY  
1150 CGTGAATCAAGGAATGACCACTATCAAAATTCGCTCCACATGCTGAGAGGCTGGAGAT 1209  
Db |||||  
4873 AGAAGGAACGAGGCATGATAGCTTTCTGTTCCGACCTCACTGTGGAGAAGCTGGAGCC 4932  
QY  
1210 GTTGATCACTTGGCAGGACATTTCTCTCTGTCACAACATATCAATGAAATTAATCTA 1269  
Db |||||  
4933 CTCACCCATCTCATGACATTCATGATGATGATGATGATGATGATGATGATGATGATGAT 4992  
QY  
1270 AGAAGTCTCTGCTGCTCAGTACTTGTACTATCTTTGGTGCAGATGCTGCGAGTCTCC 1329  
Db |||||  
4993 AAAAGAGTCCGCTGCTACAGTACTTGTGTTTCTTAGCCCAATTCCTCCGCAATGCA 5052  
QY  
1330 CATTGAGCAACAACTCTTATTTCTGACTATCATCGCAACCCCTTTTCAAGTTCCTTC 1389  
Db |||||  
5053 CCATAAGTAACAATAGCCTTATTTCTAGAGTATGCCAAAATCCCTTTTGGATTTTCTT 5112  
QY  
1390 CAACGAGTCTGAATGTCATATCTACGATGACCTTTGCAAAATTCACCTGCAAAA 1449  
Db |||||  
5113 CAGAAAGGCTAATGATCTCACTGTCTACAGATGACCCCAATGCAATTCACCTTACCAAG 5172  
QY  
1450 GAACCAATGTTGGTGAAGAATACAGCATTTGCTGCTGCTGGAAGCTCAGTTCTTGTGAT 1509  
Db |||||  
5173 GAGCCCTTAATGAAGAATATGCTATTGCTGCAAGTCTTCAAGCTGAGCACTGTGAT 5232  
QY  
1510 TTATGGAATTCGAGGAATCTGTTTACCAATCTGGGTTTTCATGCTCTCAGGCG 1569  
Db |||||  
5233 ATGTGCGAAGTGGCAAGAACAGTGTCTTGCAGTGTGGAATTTCTCATGAGGAGAAGTA 5292  
QY  
1570 CACTGGATTGGTAAGAACTACTTCAAAGAGGACCTGCTGGAATGATATTTCAGAAAC 1629  
Db |||||  
5293 AAGTTTCTGGCGACAAATTAACCTTGGAGAGGCGCTTGTGGAATGATATCCGGAGACA 5352  
QY  
1630 AATGTACCGCACATCAGGGTTCAATTTAGAGAGATGATCTGGAGAAATGAATGAACATA 1689  
Db |||||  
5353 AATGTAGCCCAATCCGATGGCCTATCGCTATGAAACCTGGTGTATGAATCAATTTA 5412  
QY  
1690 GT 1691  
Db 5413 AT 5414

RESULT 15  
AAS66419  
ID AAS66419 standard; cDNA; 5442 BP.  
XX  
AC AAS66419;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #2223.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

food supplement; medical imaging; diagnostic; genetic disorder; ss.  
Homo sapiens.  
WO200175067-A2.  
11-OCT-2001.  
30-MAR-2001; 2001WO-US08631.  
31-MAR-2000; 2000US-0540217.  
23-AUG-2000; 2000US-0649167.  
(HYSE-) HYSEQ INC.  
Dmanac RT, Liu C, Tang YT;  
WPI; 2001-639362/73.  
P-PSDB; ABG02232.  
New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity -  
Claim 1; SEQ ID No 2223; 103pp; English.  
The invention relates to isolated polynucleotide (I) and  
polypeptide (II) sequences. (I) is useful as hybridisation probes,  
polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
and gene mapping, and in recombinant production of (II). The  
polynucleotides are also used in diagnostics as expressed sequence tags  
for identifying expressed genes. (I) is useful in gene therapy techniques  
to restore normal activity of (II) or to treat disease states involving  
(II). (II) is useful for generating antibodies against it, detecting or  
quantifying a polypeptide in tissue, as molecular weight markers and as  
a food supplement. (II) and its binding partners are useful in medical  
imaging of sites expressing (II). (I) and (II) are useful for treating  
disorders involving aberrant protein expression or biological activity.  
The polypeptide and polynucleotide sequences have applications in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits to assess biodiversity  
and to produce other types of data and products dependent on DNA and  
amino acid sequences. AAS64197-AAS94564 represent novel human  
diagnostic coding sequences of the invention.  
Note: The sequence data for this patent did not appear in the printed  
specification, but was obtained in electronic format directly from WIPO  
at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 5442 BP; 1626 A; 1147 C; 1281 G; 1388 T; 0 other;  
Query Match 24.0%; Score 458.8; DB 23; Length 5442;  
Best Local Similarity 57.1%; Pred. No. 3e-75;  
Matches 857; Conservative 0; Mismatches 642; Indels 3; Gaps 1;  
QY 193 ACCTCTTCTCAGTCTGCTGATTTATTTCTCCAGTAAGCTGCTGAGGAAACACAGAACT 252  
Db |||||  
3913 ACCTTCTTAGACGATATGAAATTTTACTTGTCTTAATGCTCAAGGACCTGTAAGACC 3972  
QY  
253 GTCTGCCATAATCGGTTAAATCTTCTGAGCAATAGTTTAAATTCATCTGATGTTAAAT 312  
Db |||||  
3973 TATACCCACCGCGCTGAAGTTCCTCTCCTCAAGTTCAGGTCATCAGATGCTTAAC 4032  
QY  
313 GCGGATAGGAATTTCTTCCCGAGAGACTGCCACATCGGTGATTTTACATGTCAGS 372  
Db |||||  
4033 GAGATGGACGAGTTAAAGAGAGCTGAAACAAACCCCGGAGATTTTATACCTGACG 4092  
QY  
373 AAGGTTGACACTCATGTTTCATCATTCAGCATCATGAATCAAAAACATCTGTTGAGGTT 432  
Db |||||  
4093 AAGGTGGACACCCATATCCATCGAGCGCTTGCATGACCAACAGAACATCTGCTGCTTT 4152  
QY  
433 ATAAATCCAACTAAGAAAGAACTGATGAGGTGGTTCATTTTCAGAGATGCTATTAT 492  
Db |||||

Db 4153 ATTAGAATCTTACCAATTTGATGCTGACAGAGTGGTCTATAGACCAAAAGAGAGAT 4212  
Qy 493 ATGACTTTAAAGAGAGTTTTCAGAGCTTGGACTTAACCTGAGTATGATCTGAATGTGAT 552  
Db 4213 CTGACCCCTAAGGAACTTTTGTCTAAATTAATAATGATCCTTATGACCTGACTGTGAT 4272  
Qy 553 TTGCTAGATGCTCAGACAAAGCAGACATTTTCATCGTTTTCAGACAAATTCATCTAAA 612  
Db 4273 TCTTGGATGTTTCATGCTCGACGCCAGACCTTCCAGCGTTTTCATAGAGTTCAATGACAAA 4332  
Qy 613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAAACAGATAATCTTATT 672  
Db 4333 TATAATCTGTAGGAGCAGTACGGACCTCTACTTGAAGACAGACATTTACATT 4392  
Qy 673 CAAAGCCGTTTCTGCTGAGTGAACAAAGCAAGTTTCTCTGACCTTTCTCTAGCAAA 732  
Db 4393 AATGGGAATATTTTGGCACTATCATCAAGGAGGTAGTGGCGGACCTGGTGGAGCCAA 4452  
Qy 733 TATCAGATGGCAGATATAGGATTTCAATCTACGGAAGGAAACAGAGTGAATGGAGCCNA 792  
Db 4453 TACCAGCATGCTGAGCCCGCCGCTGCTCATATGCGCCAGTCTGATGAGTGGAGCAAA 4512  
Qy 793 CTGCAAGTTGATAGTGAACAAATGAATTCACAGTGGAAATGTTGCTGCGCTGTTGAG 852  
Db 4513 CTCTCTCTGTTGCTGCAATGCGATCCACTGCCCCAATGACATGGATGATCCAG 4572  
Qy 853 ATTCCACGCTTATATAATGTGTAAGGAAATGGGTATCGTTACATCATTTCCAAAATCTT 912  
Db 4573 GTTCCAGGATCTATGATGTTCCGTTCCAAAGATTTCTTCCACATTTTGGAAAATG 4632  
Qy 913 CTGCAACATTTTCTGTTCTCTTTTGGAGTTACTATTTGATCCAGCTTCACACCCAG 972  
Db 4633 CTGGAGATATTTTTCATGCGAGTGTGAGGCGCACCATCAACCCCGAGCTGACCCAGAA 4692  
Qy 973 CTCCATGCTCTCTGAGCAGAGTGTAGGTTGGACCTGTTGATGATGAAGTAACCA 1032  
Db 4693 CTCAGTGTCTCTCAAGCATATCACTGGCTTGGAGTGGATGATGATGCCAACAC 4752  
Qy 1033 GAAAGC---GTCCAAAGCACATGCCACACCTGAAACAGTGGACCAATGTGTTCAAC 1089  
Db 4753 AGTGGCCACATGTTCTCTCCAGAGTCCCAAGCCCCAGAGTGGACATTTGAAAAGAT 4812  
Qy 1090 CTGCAATTTTCAATATGCGTACTGCTATGCTATTAATTCACCTTAAACAGCTG 1149  
Db 4813 CCATCTTACACTTACTATGCTTACTACATGATGTAACCAATCATGCTGCTCAACAGCTG 4872  
Qy 1150 CGTGAGTCAAGGAAATGACCACTATCAAAATTCGTTCCATGCTGGAGAGCTGGAGAT 1209  
Db 4873 AGAAAGGAACGAGGATGAATACGTTTCTGTTCCGACCTCACTGTTGGAGAGCTGGAGCC 4932  
Qy 1210 GTTGATCACTTGGCAGCAGATTTCTTCTCTGCAACATATCATGGAATTAATCTA 1269  
Db 4933 CTCACCCATCTCATGACAGCATTCATGATGAGATGATCTCTCATGGCCCTAAATTTA 4992  
Qy 1270 AGGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329  
Db 4993 AAAAAAGTCCCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5052  
Qy 1330 CCATTGAGCAACATCTCTTATTTCTGCTATCATCGCAACCTTTTCCACAGTCTCTTC 1389  
Db 5053 CCACCTAAGTAACTAGGCTATTTCTAGATATGCAAAATCTCTTTTGGATTTCTCTT 5112  
Qy 1390 CAAAGAGTCTGAATGTCTCATTATCTACGATGACCCCTTTTGCAAAATTCACCTGACAAA 1449  
Db 5113 CAGAAAGGCTAATGATCTCACTGTCTACAGATGACCCAAATGCAATTCACCTTTACCAAG 5172  
Qy 1450 GAACCATTTGGGAAAGATACAGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1509  
Db 5173 GAGCCCTTAATGGAAGAATATGCTATGCTGCAAGTCTTCAAGCTGAGCACCTGTGAT 5232  
Qy 1510 TTATGCGAAATTCGAGGAGTCTGTTTACCAATCTGGGTTTTCATGCTCTCAAGCG 1569  
Db 5233 ATGTCGAAGTGGCAAGAACAGTGTCTGTCAGTGTGAAATTTCTCATGAGGAGAAAGTA 5292

Qy 1570 CACTGATTTGTAAGAACTACTTCAAAGAGGACCTGCTGAAAATGATATTTACAGAAC 1629  
Db 5293 AAGTTTCTGGCGCAATTTACCTTGAGGAGGCCCTGCTGAAAATGATATCCGGAGGACA 5352  
Qy 1630 AATGTACCGCACATCAGGTTCAATTTAGAGAGATGATCTGGAGAAATGAAATGAACTA 1689  
Db 5353 AATGTAGCCCAATCCGCATCGCTATCGCTATGAAACCTGGTGTATGAACTCAATTTA 5412  
Qy 1690 GT 1691  
Db 5413 AT 5414

Search completed: December 15, 2003, 18:10:04  
Job time : 525 secs





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 18:01:36 ; Search time 138 Seconds  
(without alignments)  
6108.998 Million cell updates/sec

Title: US-10-019-633-1  
Perfect score: 1910  
Sequence: 1 ccaacgctccgggtgtctcc.....aaaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	79.6	4.2	1248	4	US-09-489-847-101
2	77.2	4.0	1701	4	US-09-996-243-114
3	76.2	4.0	2806	4	US-09-653-839-9
4	74.6	3.9	1641	1	US-08-300-903A-8
5	74.6	3.9	1641	4	US-08-988-197-8
6	74.4	3.9	1927	4	US-09-336-536-66
7	73.6	3.9	2550	6	5258287-23
8	73.4	3.8	1664	4	US-09-904-615-66
9	73.4	3.8	1864	4	US-09-149-476-130
10	73.2	3.8	1674	4	US-09-996-243-300
11	72.8	3.8	1098	3	US-09-248-335-35
12	72.8	3.8	1114	4	US-09-152-060-41
13	72.4	3.8	1342	4	US-09-489-847-89
14	72.2	3.8	2447	2	US-09-014-969-14
15	72	3.8	1882	3	US-09-370-253-1
16	71.6	3.7	1069	4	US-09-372-422A-7
17	71.4	3.7	441	4	US-09-601-537-10
18	71.4	3.7	4121	4	US-09-601-537-9
19	71	3.7	931	4	US-09-482-273-31
20	71	3.7	3715	4	US-09-234-245-1
21	70.8	3.7	578	4	US-09-602-877A-95
22	70.8	3.7	1602	1	US-08-530-950-3
23	70.8	3.7	1602	3	US-08-888-429A-3
24	70.8	3.7	1602	3	US-09-149-879-3
25	70.8	3.7	1602	3	US-09-057-009-3
26	70.8	3.7	3238	3	US-08-123-934A-5
27	70.8	3.7	3238	5	PCT-US94-10080-5

28	70.8	3.7	8100	4	US-09-554-337-4	Sequence 4, Appli
29	70.8	3.7	11517	1	US-07-920-281C-1	Sequence 1, Appli
30	70.8	3.7	11517	3	US-08-466-277-1	Sequence 1, Appli
31	70.8	3.7	15538	4	US-09-554-337-1	Sequence 1, Appli
32	70.6	3.7	958	2	US-08-757-046A-5	Sequence 5, Appli
33	70.6	3.7	958	3	US-09-447-208-5	Sequence 5, Appli
34	70.6	3.7	958	3	US-09-135-988-5	Sequence 5, Appli
35	70.6	3.7	958	3	US-09-277-716-5	Sequence 5, Appli
36	70.6	3.7	958	3	US-08-597-274A-5	Sequence 5, Appli
37	70.6	3.7	958	4	US-08-908-909-5	Sequence 5, Appli
38	70.6	3.7	958	4	US-09-609-161B-5	Sequence 5, Appli
39	70.6	3.7	958	4	US-08-990-103-5	Sequence 5, Appli
40	70.6	3.7	1493	1	US-08-340-820-24	Sequence 24, Appli
41	70.6	3.7	1493	1	US-08-593-535-24	Sequence 24, Appli
42	70.6	3.7	1878	4	US-09-465-558-39	Sequence 39, Appli
43	70.6	3.7	2434	4	US-09-489-847-67	Sequence 67, Appli
44	70.4	3.7	569	4	US-09-461-325-44	Sequence 44, Appli
45	70	3.7	639	4	US-09-482-273-49	Sequence 49, Appli

ALIGNMENTS

RESULT 1  
US-09-489-847-101  
; Sequence 101, Application US/09489847  
; Patent No. 6476195  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al  
; TITLE OF INVENTION: 98 Human Secreted Proteins  
; FILE REFERENCE: P2031P1  
; CURRENT APPLICATION NUMBER: US/09/489,847  
; CURRENT FILING DATE: 2000-01-24  
; EARLIER APPLICATION NUMBER: PCT/US99/17130  
; EARLIER FILING DATE: 1999-07-29  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; EARLIER APPLICATION NUMBER: 60/095,486  
; EARLIER FILING DATE: 1998-08-05  
; EARLIER APPLICATION NUMBER: 60/096,319  
; EARLIER FILING DATE: 1998-08-12  
; EARLIER APPLICATION NUMBER: 60/095,454  
; EARLIER FILING DATE: 1998-08-06  
; EARLIER APPLICATION NUMBER: 60/095,455  
; EARLIER FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 376  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 101  
; LENGTH: 1248  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-489-847-101

Query Match 4.2%; Score 79.6; DB 4; Length 1248;  
Best Local Similarity 66.1%; Pred. No. 1.6e-09;  
Matches 115; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
Qy 1737 AGATGTCAGCTCGTGTATACACGAGTGGTGTGTAGCTGCTATGGGAATTACT 1796  
Db 1074 AGATTTTGGCTCTCTGTATGTGACACATTTTGTACATTATGTTATGGAAITTTCTT 1133  
Qy 1797 TCATGTTTGGTATGCTTTTCCTTATCTATGTCGAAATTCAACTTCGAACTTCAAAAAA 1856  
Db 1134 TCATACATTATTCCTCTAAACTCTCAAAAAAATAAAAAAAAAAAAAAAAAA 1193  
Qy 1857 AAAAAAATAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAA 1910  
Db 1194 AAAAAAATAAAAAAAAAAAAAAAAAAATAAAAAAAAAAAAAAAAAA 1247

RESULT 2  
US-09-996-243-114  
; Sequence 114, Application US/09996243

Patent No. 6478825  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC13  
CURRENT APPLICATION NUMBER: US/09/996,243  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429

;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090576  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 4.0%; Score 77.2; DB 4; Length 1701;  
Best Local Similarity 72.5%; Pred. No. 6.5e-09;  
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
QY 1773 TGTACTGCTATGGCAATTTACTTCATGTTTGGTATGCTTCTCTATCTATGCAAT 1832  
Db 1554 TTTTGCTAATTTGTTACTTTTCTTGCTAATTTGGAAGATTAACATTTTATAAAA 1613  
QY 1833 TCACTTCGAATTCACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1892  
Db 1614 TTATGCTAAGATTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1673  
QY 1893 AAAAAAATAAAAAA 1910  
Db 1674 AAAAAAATAAAAAA 1691

RESULT 3  
US-09-653-839-9

;; Sequence 9, Application US/09653839  
;; Patent No. 6433153  
;; GENERAL INFORMATION:  
;; APPLICANT: Donoho, Gregory  
;; APPLICANT: Turner, C. Alexander Jr.  
;; APPLICANT: Nehls, Michael  
;; APPLICANT: Friedrich, Glenn  
;; APPLICANT: Zambrowicz, Brian  
;; APPLICANT: Sands, Arthur T.  
;; TITLE OF INVENTION: No. 6433153el Human Calcium Dependent Proteases  
;; TITLE OF INVENTION: and Polynucleotides Encoding the Same  
;; FILE REFERENCE: LEX-0038-USA  
;; CURRENT APPLICATION NUMBER: US/09/653,839  
;; PRIOR FILING DATE: 2000-09-01  
;; PRIOR APPLICATION NUMBER: US 60/152,057  
;; PRIOR FILING DATE: 1999-09-02  
;; NUMBER OF SEQ ID NOS: 9  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO. 9  
;; LENGTH: 2806  
;; TYPE: DNA  
;; ORGANISM: homo sapiens  
US-09-653-839-9

Query Match 4.0%; Score 76.2; DB 4; Length 2806;  
Best Local Similarity 76.9%; Pred. No. 1.3e-08;  
Matches 93; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
QY 1790 TTATACTTCATGTTTGGTATGCTTCTCTATCTATGCAATTCGAATTCGAATTCAA 1849  
Db 2661 TTTTCTCCCATCTGTGGATACATTCTAATAATAGCACATGCCATTGGCAAAAAA 2720  
QY 1850 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1909  
Db 2721 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 2780  
QY 1910 A 1910  
Db 2781 A 2781

RESULT 4  
US-08-300-903A-8  
;; Sequence 8, Application US/08300903A  
;; Patent No. 5591630  
;; GENERAL INFORMATION:  
;; APPLICANT: Anderson, Dirk M  
;; APPLICANT: Gird, Judith G  
;; TITLE OF INVENTION: Interleukin-15 Receptors  
;; NUMBER OF SEQUENCES: 15  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Immunex Corporation  
;; STREET: 51 University Street  
;; CITY: Seattle  
;; STATE: Washington  
;; COUNTRY: USA  
;; ZIP: 98101  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: Apple Macintosh  
;; OPERATING SYSTEM: Apple Operating System 7.1  
;; SOFTWARE: Microsoft Word for Apple, Version 5.1a  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/300,903A  
;; FILING DATE: 06-SEPTEMBER-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: USSN 08/236,919  
;; FILING DATE: 06-MAY-1994  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Perkins, Patricia Anne  
;; REGISTRATION NUMBER: 34,695

```

; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..839
; US-08-300-903A-8

Query Match 3.9%; Score 74.6; DB 1; Length 1641;
Best Local Similarity 85.6%; Pred. No. 2.6e-08;
Matches 83; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 1814 TTCTCTATCTATGGCAAAATTCACCTTCGAACTTCACAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1873
DB 1516 TACCTTGTTATAACAAATAATAAACATCTATTTCAATA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1575

QY 1874 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910
DB 1576 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA 1612

RESULT 5
US-08-988-197-8
; Sequence 8, Application US/08988197
; Patent No. 6548065
; GENERAL INFORMATION:
; APPLICANT: Anderson, Dirk M
; APPLICANT: Giri, Judith G
; TITLE OF INVENTION: Interleukin-15 Receptors
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Immunex Corporation
; STREET: 51 University Street
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Apple Operating System 7.1
; SOFTWARE: Microsoft Word for Apple, Version 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/988,197
; FILING DATE:
; CLASSIFICATION: 121097
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/300,903
; FILING DATE: 06-SEPTEMBER-1994
; APPLICATION NUMBER: USSN 08/236,919
; FILING DATE: 06-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2822-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-587-0430
; TELEFAX: 206-233-0644
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1641 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

```
; SEQ ID NO:23:
; LENGTH: 2550
5258287-23

Query Match          3.9%; Score 73.6; DB 6; Length 2550;
Best Local Similarity 69.4%; Pred. No. 5.1e-08;
Matches 100; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1767 TTGCGTTGAGTGTGATGGGAATATATCTTCATGTTTGGTGATGCTTTCCTTATCTATG 1826
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2395 TTAACCTTTTGGGGGAAAGATATTTTGAGAAGTTTGCTTGCATGTATTTATAAATA 2454

QY 1827 GCAAATTCACCTCGAATTCACAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1886
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2455 GTAATAAAGTTTACCATTAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2514

QY 1887 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2515 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2538

RESULT 8
US-09-904-615-66
; Sequence 66, Application US/09904615
; Patent No. 6566325
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 49 Human Secreted Proteins
; FILE REFERENCE: P2032P1
; CURRENT APPLICATION NUMBER: US/09/904,615
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: 09/511,554
; PRIOR FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 60/097,917
; PRIOR FILING DATE: 1998-08-25
; PRIOR APPLICATION NUMBER: 60/098,634
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 170
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 66
; LENGTH: 664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (63)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-66

Query Match          3.8%; Score 73.4; DB 4; Length 664;
Best Local Similarity 68.7%; Pred. No. 3.7e-08;
Matches 101; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1764 GAGTTGCGTTGAGTGTGATGGGAATATATCTTCATGTTTGGTGATGCTTTCCTTATCT 1823
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 GAGGTTTTTGCTTTTTTTTTTTTTTTTTTTTGGTATATGCTGTATCTTAACT 572

QY 1824 ATGGCAATTCACCTCGAATTCACAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1883
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 573 TCCAGAAATAACGTTATATGGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 632

QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 633 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 659

RESULT 9
US-09-149-476-130
; Sequence 130, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
```





Patent No. 6478825  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi J.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Fong, Sherman  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Auscin L.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2730PIC13.  
CURRENT APPLICATION NUMBER: US/09/996,243  
CURRENT FILING DATE: 2001-11-14  
PRIOR APPLICATION NUMBER: 60/049787  
PRIOR FILING DATE: 1997-06-16  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/075945  
PRIOR FILING DATE: 1998-02-25  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/087106  
PRIOR FILING DATE: 1998-05-28  
PRIOR APPLICATION NUMBER: 60/087607  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087609  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087759  
PRIOR FILING DATE: 1998-06-02  
PRIOR APPLICATION NUMBER: 60/087827  
PRIOR FILING DATE: 1998-06-03  
PRIOR APPLICATION NUMBER: 60/088021  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088025  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088026  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088028  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088029  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088030  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088033  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088326  
PRIOR FILING DATE: 1998-06-04  
PRIOR APPLICATION NUMBER: 60/088167  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088202  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088212  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088217  
PRIOR FILING DATE: 1998-06-05  
PRIOR APPLICATION NUMBER: 60/088655  
PRIOR FILING DATE: 1998-06-09  
PRIOR APPLICATION NUMBER: 60/088734  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088738  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088742  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088810  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088824  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088826  
PRIOR FILING DATE: 1998-06-10  
PRIOR APPLICATION NUMBER: 60/088858  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088861  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/088876  
PRIOR FILING DATE: 1998-06-11  
PRIOR APPLICATION NUMBER: 60/089105  
PRIOR FILING DATE: 1998-06-12  
PRIOR APPLICATION NUMBER: 60/089440  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089512  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089514  
PRIOR FILING DATE: 1998-06-16  
PRIOR APPLICATION NUMBER: 60/089532  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089538  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089598  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089599  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089600  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089653  
PRIOR FILING DATE: 1998-06-17  
PRIOR APPLICATION NUMBER: 60/089801  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089907  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089908  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: 60/089947  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089948  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/089952  
PRIOR FILING DATE: 1998-06-19  
PRIOR APPLICATION NUMBER: 60/090246  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090252  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090254  
PRIOR FILING DATE: 1998-06-22  
PRIOR APPLICATION NUMBER: 60/090349  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090355  
PRIOR FILING DATE: 1998-06-23  
PRIOR APPLICATION NUMBER: 60/090429

; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090431  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090435  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090444  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090445  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090472  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090535  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090540  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090542  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090557  
; PRIOR FILING DATE: 1998-06-24  
; PRIOR APPLICATION NUMBER: 60/090676  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090678  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090690  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090694  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090695  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090696  
; PRIOR FILING DATE: 1998-06-25  
; PRIOR APPLICATION NUMBER: 60/090862  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/090863  
; PRIOR FILING DATE: 1998-06-26  
; PRIOR APPLICATION NUMBER: 60/091360  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091478  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091544  
; PRIOR FILING DATE: 1998-07-01  
; PRIOR APPLICATION NUMBER: 60/091519  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091626  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091633  
; PRIOR FILING DATE: 1998-07-02  
; PRIOR APPLICATION NUMBER: 60/091978  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/091982  
; PRIOR FILING DATE: 1998-07-07  
; PRIOR APPLICATION NUMBER: 60/092182  
; PRIOR FILING DATE: 1998-07-09

Query Match 3.8%; Score 73.2; DB 4; Length 1674;  
Best Local Similarity 71.6%; Pred. No. 5.5e-08;  
Matches 96; Conservative 0; Mismatches 38; Indels 0; Gaps 0;  
  
QY 1777 GCTGCTATGGAAATTAATCTTCAATGTTTGGTATGCTTTCTTCTATCTATGCGCAAAATCAA 1836  
Db |||||  
1514 GCTTCTCAGGACGAGGACATCTTGGCAGTCTGGCTTGGCCATTAAATGGAACCTGAA 1573  
QY |||||  
1837 CTTTCGAACITTCAAAAAATAA 1896  
Db |||||  
1574 GGCCTAAAAAATAA 1633  
QY |||||  
1897 AAAAAAATAAAAAAAAAA 1910  
Db |||||  
1634 AAAAAAATAA 1647

RESULT 11  
US-09-248-335-35

; Sequence 35, Application US/09248335  
; Patent No. 6096504  
; GENERAL INFORMATION:  
; APPLICANT: MCGONIGLE, BRIAN  
; APPLICANT: O'KEEF, DANIEL  
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES  
; FILE REFERENCE: CL-1128-A  
; CURRENT APPLICATION NUMBER: US/09/248,335  
; CURRENT FILING DATE: 1999-02-10  
; EARLIER APPLICATION NUMBER: 08/924,759  
; EARLIER FILING DATE: 1997-September-05  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: Microsoft Word Version 7.0A  
; SEQ ID NO 35  
; LENGTH: 1098  
; TYPE: DNA  
; ORGANISM: maize  
US-09-248-335-35

Query Match 3.8%; Score 72.8; DB 3; Length 1098;  
Best Local Similarity 79.6%; Pred. No. 6e-08;  
Matches 86; Conservative 0; Mismatches 22; Indels 0; Gaps 0;  
  
QY 1803 TTGCTATGCTTTCTTATCTATGCAAAATTCAACTTCGAACTTCAAAAAAAAAAAAA 1862  
Db |||||  
965 TCTTTTATTGTTGTTCTTTAAACAAGAGTTATATTTTACCATCAAAAAAAAAAAAA 1024  
QY 1863 AAA 1910  
Db |||||  
1025 AAA 1072

## RESULT 12:

; Sequence 41, Application US/09152060  
; Patent No. 6448230  
; GENERAL INFORMATION:  
; APPLICANT: ROSEN ET AL.  
; TITLE OF INVENTION: 28 Human Secreted Proteins  
; FILE REFERENCE: PZ003PI.US  
; CURRENT APPLICATION NUMBER: US/09/152,060  
; CURRENT FILING DATE: 1998-09-11  
; EARLIER APPLICATION NUMBER: PCT/US98/04858  
; EARLIER FILING DATE: 1998-03-12  
; EARLIER APPLICATION NUMBER: 60/040,762  
; EARLIER FILING DATE: 1997-03-14  
; EARLIER APPLICATION NUMBER: 60/040,710  
; EARLIER FILING DATE: 1997-03-14  
; EARLIER APPLICATION NUMBER: 60/050,934  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,100  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,357  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/048,189  
; EARLIER FILING DATE: 1997-05-30  
; EARLIER APPLICATION NUMBER: 60/057,765  
; EARLIER FILING DATE: 1997-09-05  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/068,368  
; EARLIER FILING DATE: 1997-12-19  
; NUMBER OF SEQ ID NOS: 118  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 41  
; LENGTH: 1114  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-152-060-41

Query Match 3.8%; Score 72.6; DB 4; Length 1114;  
Best Local Similarity 73.2%; Pred. No. 6.7e-08;  
Matches 90; Conservative 2; Mismatches 31; Indels 0; Gaps 0;



Search completed: December 15, 2003, 21:08:09  
Job time: 142 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 20:04:30 ; Search time 605 Seconds  
(without alignments)  
10492.685 Million cell updates/sec

Title: US-10-019-633-1  
Perfect score: 1910  
Sequence: 1 ccacgcgcgggtgtctcc.....aaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2201672 seqs, 1661799599 residues

Total number of hits satisfying chosen parameters: 4403344

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028	53.8	1803	10	US-09-938-842A-1715
2	471.6	24.7	3386	13	US-09-873-319-446
3	471.6	24.7	3386	13	US-09-960-706-706
4	458.8	24.0	5442	15	US-10-119-926-83
5	422	22.1	3915	13	US-10-205-219-188
6	282	14.8	282	9	US-09-294-0938-2101
7	177.6	9.3	278	9	US-09-923-876-5583
8	145.2	7.6	500	11	US-09-991-936-72
9	87.2	4.6	438	10	US-09-960-352-3507
10	77.2	4.0	255	13	US-09-930-213-630
11	77.2	4.0	1672	13	US-10-063-735-17
12	77.2	4.0	1672	13	US-10-063-526-17
13	77.2	4.0	1672	13	US-10-063-586-17
14	77.2	4.0	1672	13	US-10-063-510-17
15	77.2	4.0	1672	13	US-10-063-514-17

16	77.2	4.0	1672	13	US-10-063-516-17	Sequence 17, Appl
17	77.2	4.0	1672	13	US-10-063-523-17	Sequence 17, Appl
18	77.2	4.0	1672	13	US-10-063-527-17	Sequence 17, Appl
19	77.2	4.0	1672	13	US-10-063-528-17	Sequence 17, Appl
20	77.2	4.0	1672	13	US-10-063-529-17	Sequence 17, Appl
21	77.2	4.0	1672	13	US-10-063-536-17	Sequence 17, Appl
22	77.2	4.0	1672	13	US-10-063-540-17	Sequence 17, Appl
23	77.2	4.0	1672	13	US-10-063-546-17	Sequence 17, Appl
24	77.2	4.0	1672	13	US-10-063-562-17	Sequence 17, Appl
25	77.2	4.0	1672	13	US-10-063-564-17	Sequence 17, Appl
26	77.2	4.0	1672	13	US-10-063-565-17	Sequence 17, Appl
27	77.2	4.0	1672	13	US-10-063-568-17	Sequence 17, Appl
28	77.2	4.0	1672	13	US-10-063-570-17	Sequence 17, Appl
29	77.2	4.0	1672	13	US-10-063-577-17	Sequence 17, Appl
30	77.2	4.0	1672	13	US-10-063-579-17	Sequence 17, Appl
31	77.2	4.0	1672	13	US-10-063-581-17	Sequence 17, Appl
32	77.2	4.0	1672	13	US-10-063-582-17	Sequence 17, Appl
33	77.2	4.0	1672	13	US-10-063-583-17	Sequence 17, Appl
34	77.2	4.0	1672	13	US-10-063-584-17	Sequence 17, Appl
35	77.2	4.0	1672	13	US-10-063-587-17	Sequence 17, Appl
36	77.2	4.0	1672	13	US-10-063-589-17	Sequence 17, Appl
37	77.2	4.0	1672	13	US-10-063-591-17	Sequence 17, Appl
38	77.2	4.0	1672	13	US-10-063-592-17	Sequence 17, Appl
39	77.2	4.0	1672	13	US-10-063-593-17	Sequence 17, Appl
40	77.2	4.0	1672	13	US-10-063-596-17	Sequence 17, Appl
41	77.2	4.0	1672	13	US-10-063-597-17	Sequence 17, Appl
42	77.2	4.0	1672	13	US-10-063-600-17	Sequence 17, Appl
43	77.2	4.0	1672	13	US-10-063-602-17	Sequence 17, Appl
44	77.2	4.0	1672	13	US-10-063-604-17	Sequence 17, Appl
45	77.2	4.0	1672	13	US-10-063-606-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-09-938-842A-1715  
; Sequence 1715, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1715  
; LENGTH: 1803  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1715

Query Match 53.8%; Score 1028; DB 10; Length 1803;  
Best Local Similarity 76.2%; Pred No. 3.6e-231;  
Matches 1265; Conservative 0; Mismatches 395; Indels 0; Gaps 0;  
QY 13 GTTGCTCCATGGGAGAGGAGGTCTAATAATGACCCCTGACTTCCAAAAACCTAACCCCAAC 72  
Db 100 GTTGCAACCATGGGAAAAAGAGTCTATCTGATCTCTGACTTCCAAAGCCTAATACAGAG 159  
QY 73 CGGTTCATTATGTCCTGAAACCAAGTCAGAGATGTTTCCAAACTGTTGATGCGGTT 132  
Db 160 CCATTTGCACACTATCTCTCAGGAAAAATCTGATCATCTGTTTGAGATCAAGATGGGTT 219

QY 133 ATCCATGTTTATGCGGATAAAGATTGTACGGAGAGCATTTATCCTGCTGCTGATGCTACA 192  
DB 220 GTCCACGCTGTTGCAAAATTAAGATGCAAAAGAGATCTCTTCCGGTAGCTGATGCCACA 279  
QY 193 ACCTCTTCACGTACTGCAATATATATCTCCGAGTAAACGGCTGCGAGGAAACACAAAGACT 252  
DB 280 GGGTTTTTCACTGACTTGCATCAGCTACTCAAGTCTATAGCTGCGAGAAACATCCGACT 339  
QY 253 GTCTGCATTAATCGGTTAAATCTTCTGAGCATAGTTTAAATTCATCTCATGATGTTAAAT 312  
DB 340 TTGTGACACCGCTGAGTACTGCTTCTAGAACAGAAATTTAATCTCCATTTGATGCTTAAT 399  
QY 313 GCGGATAGGGAATTTCTCCGAGAGACTGCCCCACATCGTGATTTTACAAATGTCAGG 372  
DB 400 GGGATTAAGAAATTTCTGCTCAAAAAGTGCACCATCGTGATTTTATACGTTAGG 459  
QY 373 AAGGTTGACACTCATGTTTCATCATTCAGCATGATGAATCAAAAACATCTGTTGAGGTTT 432  
DB 460 AAGTCGACACTCATGTCATCATTCAGCTTGCATGAACACGAAACACCTTTTAAAGTTT 519  
QY 433 ATAAATCCAACTAAGAAAGAACCTGATGAGTGGTCTATTTTCAGAGATGCTACTTAT 492  
DB 520 ATTAAGTCAAGCTCCGAAAGAACCCGATGAGGTTGTAATATTCGAGATGGAACATAT 579  
QY 493 ATGACTTTAAAGAGGTTTTTGAGAGCTTGGACTTAACTGGGTATGATCTGAATGTTGAT 552  
DB 580 TTGACCTTGAGAGAAAGTTTTGAGACCTGGATCTGACTGGATATGACCTGAAAGTCGAC 639  
QY 553 TTGCTAGATGTCATGACAGACAAAAGACATTTTCATCGTTTTCGACAAATTCATCTAAA 612  
DB 640 CTTTGTGATGTTTCATGCAGACAAAAGTACTCTTTCATCGTTTTCGATAGTTTCAACCTAAAG 699  
QY 613 TACAATCCATGTCGCAAGTAGGCTCAGAGAAATTTTCTCTCAAAAGATATCTTAT 672  
DB 700 TATAACCTTGTGTCGCAAGTAGGCTTAGGAGATTTTCTTAAACAGGATAATCTCATC 759  
QY 673 CAAGGCGGTTTTCTCTGAGTTGACAAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAA 732  
DB 760 CAAGGTCGATTTCTTGAGATTAACAAGCAAGTCTTCTGACCTTGAAGCTAGTAA 819  
QY 733 TATCAGATGCGCAATATATAGGATTTTCAATCTAGGAAAGGAAACAGAGTGAATGGGACCAA 792  
DB 820 TATCAGATGCTGAATACAGATATCTATATATGCGAGAAATAGAGAGTGGGACCAA 879  
QY 793 CTTGCAAGTTGATGATGAATGAATTCAGCAGTGGAAATGTTGCTGGCTGGTTGAG 852  
DB 880 CTGCTAGTTGGATTTGACCAATGATCTATACAGTGAAGATGTTGCTGCTGTTAATTCAG 939  
QY 853 ATTCCACGCTTATATATATGTTACAAGGAAATGGGTATCGTTACATCATTTCCAAATCTT 912  
DB 940 CTCCACGCTTGTACAAATTTACAAGGACATGGGTAATTTGACATCGTTCCAGATATC 999  
QY 913 CTTGACAAATTTTCGTTCTCTTTTGGAGTTTACTATTGATCCAGCTTTCACACCCACAG 972  
DB 1000 CTGGACAATATATTCATCTCTGTTTGAAGCAGCTAGATCTGATTTCCCATCTCTCAG 1059  
QY 973 CTCCATGCTTCTGAGCAGGTTTGGAGTTGGACCTGGTTGATGATGAATGAAGTAAACCA 1032  
DB 1060 CTCCATGTTTTTTTGAAGCAGGTTTGGATTTGATTTGATTTGATGATGAAGAACCACT 1119  
QY 1033 GAAAGCGCTCCAAAGACATGCCACACCTGAAACAGTGGACCAATGTTTCAACCT 1092  
DB 1120 GAAAGAGCTCCCAAAACACATGCCCATCTCCAGCTCAATGGAATCAACGATCAATCTC 1179  
QY 1093 GCATTTTCATATATGCTACTACTGCTATGCTAACTATTTCACCTTAAACAGCTGCGT 1152  
DB 1180 GCATTTTCGATATATGCTACTATTGTTATGCTAACTCTATGTTGTTAAATAAGCTTGA 1239  
QY 1153 GAGTCAAGGGAATGACCACTATCAAAATTCGCTCCACATGCTGGAGAGGCTGAGATGTT 1212  
DB 1240 GAGTCAAGGGCATGACTACTATACGCTACGACCACTTCTGGAGAGGCTGGTGACATT 1299

QY 1213 GATCATTGGCAGCAGACATTTCTCTGTGTCACAAACATATCACATGGAATTAATCAAGG 1272  
DB 1300 GACCACTTGGCTGCTACGTTTCTTAACATGCCATAGCATCGCACATGGAATCAATCTGCGA 1359  
QY 1273 AAGTCTCTGCTGCTTCACTATCTTGTATCTTGTGTCAGATTTGCTGGCGATGTCGCCA 1332  
DB 1360 AAGTCTCTGCTGCTTCACTATCTGATCTACCTCGCCAGATTTGGTCTGGCCATGTCACCA 1419  
QY 1333 TTGACACAACTCCTTATTTCTGACTATCATCGCAACCTTTTCCAAAGCTTCTTCCAA 1392  
DB 1420 CTGAGCAACAACTCTTTGTTCTAGATTACCACCGGAACCCGTTTCTCTGTGTTTTCTTA 1479  
QY 1393 CGAGTCTGAATGCTCTCAATATCTACGATGACCCCTTTGCAAAATTCACCTGACAAAAGAA 1452  
DB 1480 AGAGTCTCAATGTTTCTCTGCTACTGATGACCCCTTCAGATTCATTAACATAAGAA 1539  
QY 1453 CCATTTGGGAAGAAATACAGCATTTGCTGCTGCTGGTGGGAAGCTCAGTTCTTGTGATTTA 1512  
DB 1540 CTTCTCGTGAAGAGTATAGCATAGCTGCATCAGTTTGGAAAGCTGAGTGCCTGTGACCTG 1599  
QY 1513 TCGGAAATTCGAGAACTCTGTTTACCAATCTGGGTTTTTCCATGCTCTCAAGGCGCAC 1572  
DB 1600 TCGGAGATAGCTCGTACTCAGTGTACAGTCAAGTTTCTCACAGCCCTGAGTGGCAC 1659  
QY 1573 TGGATTGTAAGAACTACTTCAAAGAGGAGCTGCTGGAATGATTAATTCACAGAACCAAT 1632  
DB 1660 TGGATTGGAAGATTTACTACAAAGAGGAGCTGATGGAACGACATTCACAAACAAAC 1719  
QY 1633 GTACCGCACATCAGGGTTCAATTTAGAGAGATGATCTGGA 1672  
DB 1720 GTGCCACACATAAGGGTGGAGTTCCGTGACACGGTATGGA 1759

## RESULT 2

US-09-873-319-446  
; Sequence 446, Application US/09873319A  
; Publication No. US20030134324A1  
; GENERAL INFORMATION:  
; APPLICANT: Munger, William E.  
; APPLICANT: Kulkarni, Prakash  
; APPLICANT: Getzenberg, Robert H.  
; APPLICANT: Waga, Iwao  
; APPLICANT: Yamamoto, Jun  
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic  
; TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles  
; FILE REFERENCE: 44921-5029-US  
; CURRENT APPLICATION NUMBER: US/09/873,319A  
; EARLIER FILING DATE: 2001-06-05  
; EARLIER APPLICATION NUMBER: US 60/223,323  
; NUMBER OF SEQ ID NOS: 755  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 446  
; LENGTH: 3386  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 M91029  
US-09-873-319-446

Query Match 24.7%; Score 471.6; DB 13; Length 3386;  
Best Local Similarity 58.3%; Pred. No. 2.9e-100;  
Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;  
QY 238 GGGAAACAAGAACTGCTGCAATATCGTTAAATCTTTTGAGCATAGTTTAAATTC 297  
DB 956 GGGCCCAATAAAGTCAATCTCTGTACCGCGGGTGCAGTACCTCAGCTCCAAAGTTCAGATG 1015  
QY 298 CATCTGATGTTAAATCGGATAGGAATTTCTTCCCAAGAGACTGCCCCACATCGTAT 357  
DB 1016 CATGCTGCTCAATGATGATGAAGAGCTGGCCGCCAGAGAAAGTGCCACCCGAGAT 1075  
QY 358 TTTTACAATGTGAGGAAGGTTGACACTCATGTTTCAATTCAGCATGATGAATCAAAA 417

Db 1076 TTCTACAAATCCGAGGTTGACACCCACATCCATGCTCTGCTCGATGAACCAAG 1135  
Qy 418 CATCTGTGAGGTTTCATAAAATCAAACTAAGAAAAAGAACTGATGAGGTGGTCAATTTTC 477  
Db 1136 CATCTGTGCTGCTTCATCAAGCGGCAATGAAGCGGCACCTGGAGGAGATCGTGACGG 1195  
Qy 478 AGAGATGGTACTTATATGACTTTAAAGAGGTTTTTGGAGCTGGACCTTAACCTGGGTAT 537  
Db 1196 GAGCAGGGCGGTGAACAGACGCTGCGGGAGGTCTTTGAGAGCATGAATCTCAAGGCTAC 1255  
Qy 538 GATCTGAATGTTGATTTGCTAGATGTCATGACAGAAAGACATTTTCATGCTTTTGAC 597  
Db 1256 GACTGAGTGGACACGCTGGAATGTCATCGGACAGGAACATTTTCATGCTTTGAC 1315  
Qy 598 AAATTCAAATCAAAATACAAATCATGCTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAA 657  
Db 1316 AAGTTTAATGCCAAATACAAACCTATTGGGGAGTCCGTCTCGAGAGATCTTCATCAAG 1375  
Qy 658 CAAGATAATCTTATTCAAGGCGGTTTTCTTGCTGAGTTGACAAAGCAAGTTTTCTCTGAC 717  
Db 1376 ACGGACAACAGGTAATCTGGGAAGTATTTTGCTCACATCATCAAGGAGGTGATGTCAGAC 1435  
Qy 718 CTTTCTGTAGCAAAATATCAGATGGCAGATATAGGATTTCAATCTACGGAAGAAACAG 777  
Db 1436 CTGAGGAGCAAAATACCAAGATGAGAGCTCGGCTCTCCATTTACGGGCGCTCGAGG 1495  
Qy 778 AGTGAATGGACCAACTTGAAGTTGGATGTAACAATGAATTCACACAGTGAATGTT 837  
Db 1496 GATGAGTGGACAAGCTGGCGGCTGGCGGCTCATGACCGGCTGCACTCCCAACGCTG 1555  
Qy 838 GTCTGCTGTTGAGATTCACGCTTATATGATGTAACAAGAAATGGGTATCGTTACA 897  
Db 1556 CGTGTGCTGGTGGAGTGGCGGCTCTTTGATGTCAGCTACCAAGGCGCAGCTGGCC 1615  
Qy 898 TCATTCCAAATCTTCTGACACATTTTGGTCTCTTTTGGAGTTACTATTGATCA 957  
Db 1616 AACTTCAGAGAGATCTGGAGAACATCTTCTCGCCACTGTTCAGGCGCACTGTGCACCT 1675  
Qy 958 GCTTTCACCCACAGCTCATGCTCTTCTGAAGCAGGTTGTAGGTTGACCTGGTTGAT 1017  
Db 1676 GCCAGCACCGGAACTGATCTCTTTAGAGCAGCTGGATGTTTTCAGACGCTGAT 1735  
Qy 1018 GATGAAGTAAACCAAGAGGCTCCAAACAAAGC---ACATGCCACACCTGAACAGTGG 1074  
Db 1736 GATGAGTCCAAAGCTTGAACCAATGCTTCAACCTGGAGAGCCCTGCTGAGGCGTG 1795  
Qy 1075 ACCAATGTTCAACCTGCAATTTTCATATATGCTACTGCTATGCTACTGATCACTATTTC 1134  
Db 1796 GTGGAGGAGAACCCACCTATGCTTACTACCTGTACTACACCTTTTCCCAACATGGCC 1855  
Qy 1135 ACCCTAAACAAGCTGCTGAGTCAAGGGAATGACCACTATCAAAATTCGCTCCACATGCT 1194  
Db 1856 ATGTTGAACCACTGCGCAGGAGGCTTCCACAGCTTTGCTGAGGCGCACTGTT 1915  
Qy 1195 GAGAGGCTGGAGATTTGATCATCTGGCAGGACATTTCTCTCTGTGCACACATATCA 1254  
Db 1916 GGGGAGGCTGGGCCATCACCACCTGGTGTGAGCCTTCATGCTGGCTGAGAACAATTC 1975  
Qy 1255 CATGGAATTAATCAAGGAAGTCTCTGTGCTTCACTGATCTGATCTTCTGTCAGATT 1314  
Db 1976 CAGGGCTCTTCTGGCAAGGCCCGCTGCTGAGTACTGATCTGATCTGATCTGATCTG 2035  
Qy 1315 GGTCTGGGATGTCCCAATGAGCAACAACCTCTTATTTCTTGACTATCATCGCAACCT 1374  
Db 2036 GGCATGCCATGTCTCCGCTCAGCAACAACAGGCTCTTCTCAGCTATCACCGGAATCG 2095  
Qy 1375 TTTTCAACCTTTTCAACGAGTCTGAATGTCTCATTTATCTACGAGTACCCCTTTGCAA 1434  
Db 2096 CTACCGGAGTACTGTGTCGGCGGCTCATGGTCTCCCTGTCACCTGATGATCCCTTGCAG 2155  
Qy 1435 ATTCACTCACAAGAACCATTTGTTGGAAGATACAGCATTTGCTGCTTCGCTGTGGAG 1494

Db 2156 TTCCACTTCCACCAAGGAGCCGCTGATGAGGAGTACAGATCGCCACCACCGGTGGAG 2215  
Qy 1495 CTCAGTCTTGTGATTTATGCGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTCA 1554  
Db 2216 CTCAGTCTCTCGATATGTGAGCTGGCCGCAACAGCGTCTCATGAGCGGCTTCTCG 2275  
Qy 1555 CATGCTCTCAAGGCGCACTGATTTGTAAGAACTACTCTTCAAAAGAGGACCTGCTCGAAAT 1614  
Db 2276 CACAAGGTAAGAGGCACTGCTGGCTGGAGCCCACTATACCAAGGAAGGCTGAGGGGAT 2335  
Qy 1615 GATATTTCACAGAACCAATGTACCGCACATCAGGTTTCAATTTAGAGAGATGATCTGGAGA 1674  
Db 2336 GACATCCCGGACCAATGTGCCAGACATCCCGTGGGTACCGCTACGAGACCTGTGC 2395  
Qy 1675 AATGAATGAACCT 1688  
Db 2396 CAGGAGCTGGCGCT 2409

RESULT 3  
US-09-960-706-706  
; Sequence 706, Application US/09960706  
; Publication No. US20030134280A1  
; GENERAL INFORMATION:  
; APPLICANT: Munger, William E.  
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia  
; FILE REFERENCE: 44921-5029-01US  
; CURRENT APPLICATION NUMBER: US/09/960,706  
; CURRENT FILING DATE: 2001-09-24  
; PRIOR APPLICATION NUMBER: 60/223,323  
; PRIOR FILING DATE: 2000-08-07  
; PRIOR APPLICATION NUMBER: 09/873,319  
; PRIOR FILING DATE: 2001-06-05  
; NUMBER OF SEQ ID NOS: 1124  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 706  
; LENGTH: 3386  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 M91029  
US-09-960-706-706

Query Match 24.7%; Score 471.6; DB 13; Length 3386;  
Best Local Similarity 58.3%; Pred. No. 2.9e-100;  
Matches 847; Conservative 0; Mismatches 604; Indels 3; Gaps 1;

Qy 238 GGGACACAAAGTCTGCTGCCATAATCGGTTAAATCTTTTGAGCATAAAGTTTAAATTC 297  
Db 956 GGGCCCATAAAGTCAATCTGCTACCGCGGCTGCAGTACCTGAGCTCCCAAGTTCCAGATG 1015  
Qy 298 CATCTGATGTTAAATGCGGATAGGAAATTTTCCCGCAAGACTGCCCCACATCGTAT 357  
Db 1016 CATGTGCTACTCAATGAGATGAAGGAGTGGCCGCCCAAGAAAGTGCACACCGAGAT 1075  
Qy 358 TTTTACAATGTCAGAAAGTTGACATCATCTTCATCATTTCCAGCATGATGAATCAAAAA 417  
Db 1076 TTCTACAACATCCGCAAGGTGACACCCACATCATTCGCTCGTCTGATGAACAGAG 1135  
Qy 418 CATCTGTTGAGGTTTCAAAAAATCCAAAAATGAAGAAAGAACCTGATGAGGTGGTTCATTTTC 477  
Db 1136 CATCTGCTGCTTTCATCAAGCGGCAATGAAGCGCACCTGGAGGAGATCGTGACGCTG 1195  
Qy 478 AGAGATGTTACTTATATGACTTTTAAAGAGGTTTTTGGAGCTTTGAGCTTTAACTTGGGTAT 537  
Db 1196 GAGCAGGCGCTGTAACAGACGCTCGGGAGGTCTTTTGAGAGCATGAATCTCACGGCTTAC 1255  
Qy 538 GATCTGAATGTTGATTTGCTAGATGTCATGCAACAAAAAGCACATTTTCATCGTTTGAC 597  
Db 1256 GACCTGAGTGTGACACGCTGATGTGATCGGAGAGAACACTTTTCCATCGCTTTGAC 1315  
Qy 598 AAATTCATCTAAATAATCAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAAA 657



1316 AAGTTTAAATGCCAAATACAAACCTATTGGGGAGTCGCTCCGAGAGATCTTCATCAAG 1375  
658 CAGATAAATCTTATCAAGGCGGTTTCTTGCTGAGTTGACAAAGCAAGTTTCTCTGAC 717  
1376 ACGGACAAACAGGGATATCTGGGAAGTACTTTGCTACATCATCAAGGAGGTGATGCAGAC 1435  
718 CTTTCTGTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAG 777  
1436 CTGGAGAGAGCAAAATACAGAAATGAGAGCTGGGCTCTCCATTTACGGGCGCTCGAGG 1495  
778 AGTGAATGGACCAACTTCAAGTTGGATAGTGAACAATGAATGAATTCACAGTGGAAATGTT 837  
1496 GATGAGTGGACAACTGGCGGCTGGCGGCTCATGCACCGCGTGCACTCCGCCCAACGCTG 1555  
838 GTCTGCTGTTGAGATTCAGATTCACGCTTATATATATGTTGACAGGAATGGTATCGTTACA 897  
1556 CGCTGGCTGGTGGAGTGGCGGCTCTTTGATGTGACCTGACCAAGGGGCCAGCTGGGC 1615  
898 TCATTCCAAAATCTCTTGACAACTTTTGGTCTCTTTTGGAGTTACTATTGATCCA 957  
1616 AACTTCCAGGAGATGCTGAGAACATCTTCTGCCACTGTTGAGGCCACTGTGCACCTT 1675  
958 GCTTCAACCCACAGCTCCATGCTTCTTGAAGCAGGTTGTAGGTTGACCTGCTGTGAT 1017  
1676 GCCAGCCACCGGAACCTGATCTCTTCTTAGAGCAGTGGATGGTTTGACAGCGTGGAT 1735  
1018 GATGAAGTAAACCAAGAGGGGCTCAACAAAGC---ACATGCCACACCTGAACAGTGG 1074  
1736 GATGAGTCCAAAGCTGAAACCAATGCTTTCAACCTGGAGAGGCCCTGCTGAGGCGTGG 1795  
1075 ACCAATGTGTTCAACCTGATTTTCATATTATGCTGCTACTGCTATGCTACTTATTC 1134  
1796 GTGGAGAGGACAAACCCCTAGCTACTACTGCTACTGCTACTGCTACTGCTACTGCTACTG 1855  
1135 ACCCTAAACAGCTGGCTGAGTCAAAAGGGAATGACCACTATCAAAATCCCTCCACATGCT 1194  
1856 ATGTTGAACACCTGCGCAGGAGAGGGCTTCCACACGTTTGTGCTGAGGCCACACTGT 1915  
1195 GAGAGGCTGGAGATGTGATGACTTGTGGCAGCAGATTTCTTCTGTCACACATATCA 1254  
1916 GGGGAGGCTGGGCCCCATCCACCACCTGGTGTGACGCTTCAATGCTGCTGAGAACATTC 1975  
1255 CATGGAATTAATCTAAGGAAGTCTCTGTGCTTTCAGTACTGTACTATCTGCTGCTAGATT 1314  
1976 CAGGGCTCTTCTGCGCAAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2035  
1315 GGTCTGGCGATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1374  
2036 GGCATCGCCATGCTCGCTCAGCAACACAGCCTCTTCTCAGCTATCACCGGAATCCG 2095  
1375 TTTCCAAAGTCTTCCAAAGGCTGGAATGTCTCATTTATCTACGATGACCTTTGCA 1434  
2096 CTACCGGAGTACCTGCTCGCGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2155  
1435 ATTCACCTGCAAAAGAACCAATGTTGGAAGATACAGCAATTCCTGCTGCTGCTGCTGCTG 1494  
2156 TTCCATTTACCAAGAGCGCTGATGAGAGAGTACAGCATCCGACCCAGGTGTGGAAG 2215  
1495 CTCAGTTCTTTGATTTATCGAAATTCGAGGAATCTGTTTACCAATCTGGGTTTCA 1554  
2216 CTCAGCTCTGCGATATGTGAGCTGGCCGCAACAGCGTGTCTATGAGCGGCTTCTCG 2275  
1555 CATGCTCTAAGGGGCTGATGTTGTAAGAACTCTTCAAAAGAGGACCTGCTGGAAT 1614  
2276 CACAAGGTAAAGAGCCACTGGCTGGGACCACTATACCAAGGAAGGCCCTGAGGGGAAT 2335  
1615 GATATTACAGAAACCAATGTACCGCACATCAGGGTTTCAATTTAGAGAGATGATCTGGA 1674  
2336 GACATCGCGGACCAATGTGCCAGACATCCGCTGGGCTACCGCTACGAGACCTTGTGC 2395  
1675 AATGAATGAACCT 1688

Db 2396 CAGGAGCTGGCGCT 2409  
RESULT 4  
US-10-119-926-83  
; Sequence 83, Application US/10119926  
; Publication No. US20030104413A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Zhou, Ping  
; APPLICANT: Xu, Chongjun  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030104413A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 789CIP2BCON  
; CURRENT APPLICATION NUMBER: US/10/119,926  
; CURRENT FILING DATE: 2002-04-09  
; PRIOR APPLICATION NUMBER: 09/574,454  
; PRIOR FILING DATE: 2000-05-19  
; PRIOR APPLICATION NUMBER: 09/519,705  
; PRIOR FILING DATE: 2000-03-07  
; NUMBER OF SEQ ID NOS: 108  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 83  
; LENGTH: 5442  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2251) .. (2820)  
US-10-119-926-83  
Query Match 24.0%; Score 458.8; DB 15; Length 5442;  
Best Local Similarity 57.1%; Pred. No. 3.8e-97;  
Matches 857; Conservative 0; Mismatches 642; Indels 3; Gaps 1;  
QY 193 ACCCTTCTTCACTGACTTGCATTTATTTCTCCGAGTAAACGGCTGCGAGGACACACAGAACT 252  
Db 3913 ACCCTTCTTCACTGACTTGCATTTATTTTCTGCTTAAATTTGCTCAAGGACCTGTTAAGACC 3972  
QY 253 GTCTGCCATAATCGGTTAAATCTTCTGAGCATAAGTTTAAATTCATCTGATGTTAAAT 312  
Db 3973 TATACCCACCGGCGCTGAGTTCTCTCTCCAGTTCCAGGTCCATCAGATGCTTAAC 4032  
QY 313 CGCGATAGGGAATTTCTTCCCGCAGAGACTGCCCAATCGTGATTTTACAAATGTCAGG 372  
Db 4033 GAGATGGACGAGTTAAAGAGAGCTGAAAAAACACCCGCCAGAGATTTTATAACTGCAGG 4092  
QY 373 AGGTTGACACTCATGTTTCATCATTTCCAGATGCTGAATCAAAAACATCTGTTGAGTTTC 432  
Db 4093 AAGGTGGACACCCATATCCATGACGCGCTTGCATGACGCGCTTGCATGACGCGCTTGC 4152  
QY 433 ATAAAACTCAAACTAAGAAAGAACTGATGAGGTGGTCAATTTTCAGAGATGCTACTTAT 492  
Db 4153 ATTAGAATCTTACCAATTTGATGCTGACAGAGTGGTCTATAGCAACCAAGAGAGAT 4212  
QY 493 ATGACTTTAAAAAGAGGTTTTTGAGAGCTTGGACTTAACTGGGTGATGATCTGAATGTTGAT 552  
Db 4213 CTGACCCCTAAAGGAACTTTTCTGCTAAATTTAAAAATGATCTCTTATGACCTGCTGTTGAT 4272  
QY 553 TTGCTAGATGCTCCATGCGACAGCAAAAGCAGATTTTCTGCTTTTGAACAAATTCATCTAAA 612  
Db 4273 TCTCTGGATGTTTCTGCTGGACGCGCCAGACCTTCCAGCGGTTTGTGTAAGTTCAATGACAA 4332





```

Db      243  TAAATAAGCTGATGAACAATTTCCAAAGAAATCTTACAAATATCTATTGCCATTATTG 184
Qy      941  AGGTTACTATTGATCCAGCTTTCACACCACAGCTCCATGTCTTCTCTGAAGCAGGTTGTAG 1000
Db      183  AAGTTACGAATGATCCAAATCGCATCTCTGAACCTACACAAGTTTCTTTCAGTTCGTGTAG 124
Qy      1001  GTTTGGACCTGGTTGATGATGAAGTAAACACAGAAAGCGGTCCA---ACAAAGCACATGC 1057
Db      123  GTTTTGACTCGTGGATGATGAAGTAAACCCGAAATCCATTATTTCGATAAGGATGTTTC 64
Qy      1058  CCACACTGAACAGTGGACCAATGTGTTCAACCTCGCATTTTCATATTATGCGTACTA 1115
Db      63  CGACACCAGAAGATGGACAGACACTGAAATCGCCCTATGCATATTATCAATATTA 6

RESULT 9
US-09-960-352-3507
; Sequence 3507, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTAT
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 3507
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (350)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: 16-BOVMS1-004-Q1-E1-D11
US-09-960-352-3507

Query Match 4.6%; Score 87.2; DB 10; Length 438;
Best Local Similarity 53.0%; Pred. No. 2.2e-10;
Matches 185; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

Qy      682  TTTCTTGCTGAGTGCACAAGCAAGTTTTCTCTGACCTTTCTGCTAGCAAAATATCAGATG 741
Db      4  TATTTTGCCACTATCATCAAGGAGGTAGGTGCGACACTTTGGTGGAGGCCAAGTACCAGCAT 63
Qy      742  GCAGAATATAGGATTTCAATCTACGGAAGAAACAGAGTGAATGGGACCACCTTGCAAGT 801
Db      64  GCGAGCCCGCTTGCTTCATCTATGGCGAGTCTTGAGGAGTGGAGCAAACTCTCCACC 123
Qy      802  TGGATAGTGAAACAATGAAATTGCACAGTGGAAATGTTGTCTGGCTGGTTTCAGATTCCAACG 861
Db      124  TGGTTCGTGTCAAAACCGCGTCTACTGCCCAACATGACATGATGATCCAGGTCCCGAGG 183
Qy      862  TTAATAATGTGTACAAGGAATGGGTATCGTTTACATCATTTCCAAAATCTTCTTTGACAAC 921
Db      184  ATCTATGATGTGTTCGATCTAAGAACTTCCTTCCACACTTTGGGAAGATGTTGGAGAAT 243
Qy      922  ATTTTCGTTCTCTTTTTCAGGTTTACTATTGATCCAGCTTTCACACCCACAGCTCCATGTC 981
Db      244  GTTTTCATGCGAGTGTGCAAGGCCACCATCAACCCCTCAGGCTCACCACGACCTCAGTGT 303
Qy      982  TTCTGAAGCAGGTTGTAGGTTTGGACCTGGTTGATGATGAAGAGTAAAC 1030
Db      304  TTCTCAAGCACATTACAGGCTTCGACAGTGTAGATGATGAGTCCANAC 352

RESULT 10
US-09-930-213-630/c
; Sequence 630, Application US/09930213

```

Publication No. US20030170625A1  
GENERAL INFORMATION:  
APPLICANT: ROSENTHAL, ANDRE  
APPLICANT: HINZMANN, BERND  
APPLICANT: SCHAFER, REINHARD  
APPLICANT: ZUBER, JOHANNES  
APPLICANT: TCHE-NITSE, OLEG  
APPLICANT: GRIPS, MARTIN  
APPLICANT: HELNEGEL, MARTIN  
APPLICANT: SCHMITZ, ANNE-CHANTAL  
APPLICANT: SEPS, CHRISTINE  
TITLE OF INVENTION: DETECTION OF DIFFERENTIAL GENE EXPRESSIONS  
FILE REFERENCE: ALBRE-14  
CURRENT APPLICATION NUMBER: US/09/930,213  
CURRENT FILING DATE: 2001-01-31  
PRIOR APPLICATION NUMBER: DE 10004102.7  
PRIOR FILING DATE: 2000-01-31  
NUMBER OF SEQ ID NOS: 885  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 630  
LENGTH: 255  
TYPE: DNA  
ORGANISM: Rattus sp.  
FEATURE: modified\_base  
NAME/KEY: LOCATION: (7)  
OTHER INFORMATION: a, t, c, g, other or unknown  
US-09-930-213-630

Query Match 4.0%; Score 77.2; DB 13; Length 255;  
Best Local Similarity 65.4%; Pred. No. 3.7e-08;  
Matches 106; Conservative 4; Mismatches 52; Indels 0; Gaps 0;

Qy 1749 TCGTGTATACAGACGAGTGGTGTAGCTGCTATGGGAATTATACCTTCATGTTTGGT 1808  
Db 198 TKGTTTTCCCTCCCTCTTTTATCGTTTGGGTTTATTTTCTTTTGTGCGC 139

Qy 1809 ATGCTTTCTTATCTATGCGAAATCAACTTCGAACCTCAAAAAA 1868  
Db 138 ATTATAMCCTTTTCTTTTAAATTTCCCTTTTGGGAAAAA 79

Qy 1869 AAAAAA 1910  
Db 78 AAAAAA 37

RESULT 11  
US-10-063-735-17  
Sequence 17, Application US/10063735  
Publication No. US20030138882A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3230R1C1  
CURRENT APPLICATION NUMBER: US/10/063,735  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 170  
SEQ ID NO 17  
LENGTH: 1672  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-063-735-17

Query Match 4.0%; Score 77.2; DB 13; Length 1672;  
Best Local Similarity 72.5%; Pred. No. 9.9e-08;  
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCTTATCTATGCGCAAT 1832  
Db 1525 TTTTGCTAAATTTGTTACTTTTCTTGTCTAAATTTGGAGATTAACTCATTTTAAATAAA 1584

Qy 1833 TCAACTTCGAACCTTCAAAAAA 1910  
Db 1585 TTATGCTCTAGATTAAAAA 1662

RESULT 13  
US-10-063-586-17  
Sequence 17, Application US/10063586  
Publication No. US20030176684A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
US-10-063-586-17

Query Match 4.0%; Score 77.2; DB 13; Length 1672;  
Best Local Similarity 72.5%; Pred. No. 9.9e-08;  
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCTTATCTATGCGCAAT 1832  
Db 1525 TTTTGCTAAATTTGTTACTTTTCTTGTCTAAATTTGGAGATTAACTCATTTTAAATAAA 1584

Qy 1833 TCAACTTCGAACCTTCAAAAAA 1910  
Db 1585 TTATGCTCTAGATTAAAAA 1662

RESULT 12  
US-10-063-526-17  
Sequence 17, Application US/10063526  
Publication No. US20030171550A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Watanabe, Colin K.  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3230R1C1  
CURRENT APPLICATION NUMBER: US/10/063,526  
CURRENT FILING DATE: 2002-05-02  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 170  
SEQ ID NO 17  
LENGTH: 1672  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-063-526-17

Query Match 4.0%; Score 77.2; DB 13; Length 1672;  
Best Local Similarity 72.5%; Pred. No. 9.9e-08;  
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

Qy 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCTTATCTATGCGCAAT 1832  
Db 1525 TTTTGCTAAATTTGTTACTTTTCTTGTCTAAATTTGGAGATTAACTCATTTTAAATAAA 1584

Qy 1833 TCAACTTCGAACCTTCAAAAAA 1910  
Db 1585 TTATGCTCTAGATTAAAAA 1662

RESULT 13  
US-10-063-586-17  
Sequence 17, Application US/10063586  
Publication No. US20030176684A1  
GENERAL INFORMATION:  
APPLICANT: Eaton, Dan L.  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
US-10-063-586-17

```

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,586
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-586-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db      |||||
1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAGATTAACTCACTTTTAAATAAAA 1584
QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db      |||||
1585 TTATGCTAAGTTAAATAAAA 1662
QY 1893 AAAAAA 1910
Db      |||||

RESULT 14
US-10-063-510-17
; Sequence 17, Application US/10063510
; Publication No. US20030180837A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,510
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-510-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db      |||||
1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAGATTAACTCACTTTTAAATAAAA 1584
QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db      |||||
1585 TTATGCTAAGTTAAATAAAA 1662
QY 1893 AAAAAA 1910
Db      |||||

RESULT 15
US-10-063-514-17
; Sequence 17, Application US/10063514
; Publication No. US20030181707A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,514
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-514-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db      |||||
1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAGATTAACTCACTTTTAAATAAAA 1584
QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db      |||||
1585 TTATGCTAAGTTAAATAAAA 1662
QY 1893 AAAAAA 1910
Db      |||||

Search completed: December 15, 2003, 22:56:48
Job time : 609 secs
```

```

Db      1645 AAAAAAAAAAAAAAAAAA 1662

RESULT 15
US-10-063-514-17
; Sequence 17, Application US/10063514
; Publication No. US20030181707A1
; GENERAL INFORMATION:
; APPLICANT: Eaton, Dan L.
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3230R1C1
; CURRENT APPLICATION NUMBER: US/10/063,514
; CURRENT FILING DATE: 2002-05-01
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 17
; LENGTH: 1672
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-063-514-17

Query Match      4.0%; Score 77.2; DB 13; Length 1672;
Best Local Similarity 72.5%; Pred. No. 9.9e-08;
Matches 100; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 1773 TGTAGCTGCTATGGGAATTATACCTTCATGTTTGGTATGCTTTCCTTATCTATGCGCAAT 1832
Db      |||||
1525 TTTTGCTAATTGTTACTTTTCTTTGCTAATTGGAGATTAACTCACTTTTAAATAAAA 1584
QY 1833 TCAACTTCGAACCTTCAAAAAA 1910
Db      |||||
1585 TTATGCTAAGTTAAATAAAA 1662
QY 1893 AAAAAA 1910
Db      |||||

Search completed: December 15, 2003, 22:56:48
Job time : 609 secs
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2003, Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 15, 2003, 16:43:05 ; Search time 3666 Seconds  
(without alignments)  
12662.725 Million cell updates/sec

Title: US-10-019-633-1  
Perfect score: 1910  
Sequence: 1 ccagcgctcgggtgtctcc.....aaaaaaaaaaaaaaaaaaaaa 1910

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_vrt:\*  
20: em\_gss\_fun:\*  
21: em\_gss\_mam:\*  
22: em\_gss\_mus:\*  
23: em\_gss\_pro:\*  
24: em\_gss\_rod:\*  
25: em\_gss\_phg:\*  
26: em\_gss\_vrl:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1910	100.0	1911	11	AY109001 Zea mays
2	1067.8	55.9	2196	11	AY107533 Zea mays
3	682.8	35.7	832	14	CB642483 OSJNEB02K
4	651.4	34.1	798	14	CB643461 OSJNEB04C

5	606	31.7	624	9	AW017903	AN017903 614020B05
6	590.8	30.9	877	14	CB660024	CB660024 OSJNEB01H
7	576.8	30.2	634	9	AW671750	AW671750 LG1 351.F
8	555.4	29.1	845	14	CB668055	CB668055 OSJNEB15L
9	551.2	28.9	757	14	CB671812	CB671812 OSJNEB05I
10	517	27.1	872	12	BM359276	BM359276 GA_Ea001
11	508.2	26.6	647	13	BU973129	BU973129 HB23N17K
12	507.6	26.6	676	12	BI955474	BI955474 HVSME002
13	494.6	25.9	880	10	BF268586	BF268586 GA_EB000
14	490.6	25.7	913	10	BG365787	BG365787 HVSME1000
15	488.2	25.6	909	10	BG442858	BG442858 GA_Ea001
16	487.8	25.5	3586	11	AK004759	AK004759 MJB_muscu
17	460	24.1	676	13	BQ871213	BQ871213 QG11D06.
18	456	23.9	954	10	BF268791	BF268791 GA_EB000
19	453.2	23.7	954	13	BU991366	BU991366 HD06L13R
20	445.2	23.3	658	14	CD226569	CD226569 CCC1 46 B
21	444	23.2	652	14	CA917577	CA917577 EST641724
22	443.4	23.2	461	12	BM501264	BM501264 PAC000000
23	442.2	23.2	679	13	BQ969627	BQ969627 QG14D05.
24	435.6	22.8	519	12	BI948901	BI948901 HVSME1001
25	434	22.7	622	10	AW906456	AW906456 EST342578
26	429.6	22.5	524	10	BE425619	BE425619 WHE0322 C
27	429	22.5	3402	11	AK034844	AK034844 MJB_muscu
28	415	21.7	666	12	BG886862	BG886862 EST512713
29	412	21.6	955	12	EM084714	EM084714 PPAP_10_E
30	411.2	21.5	759	10	BE034209	BE034209 MH01C02 M
31	410.8	21.5	647	13	BQ465687	BQ465687 HU04G08R
32	410.2	21.5	652	14	CB869343	CB869343 HC11A23W
33	409	21.4	803	14	CA767547	CA767547 AF53-RFf
34	406.6	21.3	893	10	BF268580	BF268580 GA_EB000
35	404.2	21.2	580	10	BE475206	BE475206 sp73e06.Y
36	403.4	21.1	566	12	BI785902	BI785902 sai30e01.
37	402	21.0	426	9	AW287505	AW287505 LG1 228 H
38	401.6	21.0	645	13	BQ994451	BQ994451 OGF7E08.Y
39	400	20.9	860	12	BM359287	BM359287 GA_Ea001
40	396.2	20.7	583	12	BG890464	BG890464 EST516315
41	396	20.7	588	10	EG648222	EG648222 EST509841
42	392.8	20.6	646	14	CA840318	CA840318 MCT036B07
43	383.4	20.1	589	10	BE362865	BE362865 DGI 90 CO
44	377.6	19.8	577	13	BQ778755	BQ778755 946114A12
45	371.4	19.4	533	9	AW831078	AW831078 sm08a08.Y

#### ALIGNMENTS

RESULT 1	AY109001	AY109001	1911 bp	mRNA	linear	HTC 17-OCT-2002
LOCUS	Zea mays PC0134657 mRNA sequence.					
DEFINITION	Zea mays PC0134657 mRNA sequence.					
ACCESSION	AY109001					
VERSION	AY109001.1 GI:21212355					
KEYWORDS	HTC.					
SOURCE	Zea mays					
ORGANISM	Zea mays					
REFERENCE	Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.					
AUTHORS	Hayney, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittsitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.					
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes					
JOURNAL	Unpublished (2002)					
REFERENCE	2 (bases 1 to 1911)					
AUTHORS	Coe, E.H.					
TITLE	Direct Submission					
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA					
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the					

maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB:

FEATURES  
source

Location/Qualifiers  
1..1911  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:637880"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Cornsensus Library"

/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

BASE COUNT 597 a 391 c 383 g 540 t

ORIGIN

Query Match 100.0%; Score 1910; DB 11; Length 1911;  
Best Local Similarity 100.0%; Pred. No. 6.9e-176;  
Matches 1910; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CCACGGCTCCGGTGTCTCCATGGGAGAGGAGTCAATAATGACCCCTGTACTCCAAA 60
DB 1 CCACGGCTCCGGTGTCTCCATGGGAGAGGAGTCAATAATGACCCCTGTACTCCAAA 60

QY 61 CTTAAACCCCAACCCGTTCACTTATGTGCTGCAACCAAGTCAGAGCATGTTTTCCAAACT 120
DB 61 CTTAAACCCCAACCCGTTCACTTATGTGCTGCAACCAAGTCAGAGCATGTTTTCCAAACT 120

QY 121 GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGACGGAGAGCATTTATCCTGTG 180
DB 121 GTTGATGGCGTTATCCATGTTTATGCGGATAAAGATTGACGGAGAGCATTTATCCTGTG 180

QY 181 GCTGATGCTPACAACTTCTTCTACTGACTTGCATTTATTTCTCCGAGTAACGGCTGCAGGG 240
DB 181 GCTGATGCTPACAACTTCTTCTACTGACTTGCATTTATTTCTCCGAGTAACGGCTGCAGGG 240

QY 241 AACACAAGAACTGTGTCATTAATCGGTTAAATCTTTTGAGCAATAAGTTTAAATTCAT 300
DB 241 AACACAAGAACTGTGTCATTAATCGGTTAAATCTTTTGAGCAATAAGTTTAAATTCAT 300

QY 301 CTGATGTTAAATCGGATAGGAAATTTCTTGCCAGAGACTGCCCCACATCGTATTT 360
DB 301 CTGATGTTAAATCGGATAGGAAATTTCTTGCCAGAGACTGCCCCACATCGTATTT 360

QY 361 TACAATGTCAGGAAGTTGCACACTCATGTTTCATTCATTGACGATGATGAATCAAAAACAT 420
DB 361 TACAATGTCAGGAAGTTGCACACTCATGTTTCATTCATTGACGATGATGAATCAAAAACAT 420

QY 421 CTGTTGAGGTTCAATAAATCCAAACTTAAGAAAGAACCTGATGAGTGGTGTCAATTTTCA 480
DB 421 CTGTTGAGGTTCAATAAATCCAAACTTAAGAAAGAACCTGATGAGTGGTGTCAATTTTCA 480

QY 481 GATGTTACTTATGACTTTAAAGAGGTTTTTGAGAGCTTGGAATTAACCTGGGTATGAT 540
DB 481 GATGTTACTTATGACTTTAAAGAGGTTTTTGAGAGCTTGGAATTAACCTGGGTATGAT 540

QY 541 CTGAATGTTGATTTGCTAGATGTCATGTCAGACAAAGACACATTTCTCGTTTTCACAAA 600
DB 541 CTGAATGTTGATTTGCTAGATGTCATGTCAGACAAAGACACATTTCTCGTTTTCACAAA 600

QY 601 TTCAATCTAAAATPACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTTCAAAACAA 660
DB 601 TTCAATCTAAAATPACAAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTTCAAAACAA 660

QY 661 GATAATCTTATTCAGAGCGGTTTTCTTGCTGATGTGACAAAGCAAGTTTTCTCTGACCTT 720
DB 661 GATAATCTTATTCAGAGCGGTTTTCTTGCTGATGTGACAAAGCAAGTTTTCTCTGACCTT 720

QY 721 TCTGCTAGCAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGT 780
```

```
DB 721 TCTGCTAGCAATATCAGATGGCAGAAATATAGGATTTCAATCTACGGAAGGAAACAGAGT 780

QY 781 GAATGGGACCAACTTGCAGTTGATAGTGAACAAATGAATTTGCACAGTGGAAATGTTGTC 840
DB 781 GAATGGGACCAACTTGCAGTTGATAGTGAACAAATGAATTTGCACAGTGGAAATGTTGTC 840

QY 841 TGGCTGGTTTCAGATTTCCACGCTTATATAATGTGTACAGGAAATGGGTATCGTTACATCA 900
DB 841 TGGCTGGTTTCAGATTTCCACGCTTATATAATGTGTACAGGAAATGGGTATCGTTACATCA 900

QY 901 TTCCAAATCTTCTTGACAACTTTCTCTCTCTTTTGAGGTACTTATGATCCAGCT 960
DB 901 TTCCAAATCTTCTTGACAACTTTCTCTCTCTTTTGAGGTACTTATGATCCAGCT 960

QY 961 TCACACCCACAGCTCCATGCTTCTGAGCAGAGTTGTAGGGTTGGACCTGGTTGATGAT 1020
DB 961 TCACACCCACAGCTCCATGCTTCTGAGCAGAGTTGTAGGGTTGGACCTGGTTGATGAT 1020

QY 1021 GAAAGTAAACCAAGAGCGTCCCAACAAAGCACATGCCACACCTGAACAGTGGACCAAT 1080
DB 1021 GAAAGTAAACCAAGAGCGTCCCAACAAAGCACATGCCACACCTGAACAGTGGACCAAT 1080

QY 1081 GTGTTTCAACCTCGCATTTTTCATATTATGCGTACTGCTATGCTAACTTATTCACCCCTA 1140
DB 1081 GTGTTTCAACCTCGCATTTTTCATATTATGCGTACTGCTATGCTAACTTATTCACCCCTA 1140

QY 1141 AACAGCTGCGTGAATCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200
DB 1141 AACAGCTGCGTGAATCAAGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAG 1200

QY 1201 GCTGAGAGTGTGATCACTTGGCAGCGACATTTCTTCTGTGCAACATATCACATGGA 1260
DB 1201 GCTGAGAGTGTGATCACTTGGCAGCGACATTTCTTCTGTGCAACATATCACATGGA 1260

QY 1261 ATTAATCTAAAGAAAGTCTCTGCTGCTTCACTTGTACTATCTTGGTCAAGTGGTCTG 1320
DB 1261 ATTAATCTAAAGAAAGTCTCTGCTGCTTCACTTGTACTATCTTGGTCAAGTGGTCTG 1320

QY 1321 GCGATGCCCCATTTGAGCAACAACTCTTATTTCTGATCATATCGCAACCCCTTTTCCA 1380
DB 1321 GCGATGCCCCATTTGAGCAACAACTCTTATTTCTGATCATATCGCAACCCCTTTTCCA 1380

QY 1381 ACCTTCTTCCAAAGAGTCTGAATGCTCATTTATCTACGATGACCCCTTTCGAAATTCAC 1440
DB 1381 ACCTTCTTCCAAAGAGTCTGAATGCTCATTTATCTACGATGACCCCTTTCGAAATTCAC 1440

QY 1441 CTGACAAAGAAACCAATTTGGTGAAGAAATPACAGCAATGCTGCTTCTGCTGGAAGCTCAGT 1500
DB 1441 CTGACAAAGAAACCAATTTGGTGAAGAAATPACAGCAATGCTGCTTCTGCTGGAAGCTCAGT 1500

QY 1501 TCTTGTGATTTATGGGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTTCATGCT 1560
DB 1501 TCTTGTGATTTATGGGAAATTCGAGGAACTCTGTTTACCAATCTGGGTTTTTTCATGCT 1560

QY 1561 CTCAGGCGCACTGGATTTGTAAGAACTACTTTCGAAAGAGGACCTCTGCGGAAATGATATT 1620
DB 1561 CTCAGGCGCACTGGATTTGTAAGAACTACTTTCGAAAGAGGACCTCTGCGGAAATGATATT 1620

QY 1621 CACGAAACCAATGTACCCGACATCAGGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAA 1680
DB 1621 CACGAAACCAATGTACCCGACATCAGGGTTCAATTTTAGAGAGATGATCTGGAGAAATGAA 1680

QY 1681 ATGAACTAGTGTACTCTGACAAATGAGATCTTAATACCAGACGAGCTGACCTGTGAAGAT 1740
DB 1681 ATGAACTAGTGTACTCTGACAAATGAGATCTTAATACCAGACGAGCTGACCTGTGAAGAT 1740

QY 1741 GTCCAGCCTCGGTATACAGACGAGTTCGCTGTAGCTGTATGGGAAATATATCTTCAT 1800
DB 1741 GTCCAGCCTCGGTATACAGACGAGTTCGCTGTAGCTGTATGGGAAATATATCTTCAT 1800

QY 1801 GTTTGGTATGCTTTCTTATCTATGGCAAAATTTCAACTTCGAACCTTCAAAAAA 1860
```

```

Db      1801 GTTTGGTATGCTTCTTCTATCTATGGCAAAATCAACTTCGAACTTCAAAAAA 1860
Qy      1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910
Db      1861 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1910

RESULT 2
LOCUS   AY107533
DEFINITION Zea mays PC0130446 mRNA sequence.
ACCESSION AY107533
VERSION   AY107533.1 GI:21210611
KEYWORDS HTC.
SOURCE   Zea mays
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 2196)
AUTHORS   Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
           Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
TITLE      Maize Mapping Project/DuPont Consensus Sequences for Design of
           Overgo Probes
JOURNAL   Unpublished (2002)
REFERENCE 2 (bases 1 to 2196)
AUTHORS   Coe,E.H.
TITLE      Direct Submission
JOURNAL   Submitted (25-APR-2002) Maize Mapping Project, University of
           Missouri, Columbia, MO 65211, USA
COMMENT   If you are interested in getting corresponding physical clones,
           these are publicly available from ZmDB and may be found by BLAST
           searching at MSL, maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR,
           www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the
           maize cDNA sequences is either Virginia Walbot, Stanford or Pat
           Schnable, Iowa State, then clones may be requested from ZmDB:
           www.zmdb.iastate.edu.

FEATURES             Location/Qualifiers
     source           1..2196
                     /organism="Zea mays"
                     /mol_type="mRNA"
                     /db_xref="MaizeDB:637602"
                     /db_xref="taxon:4577"
                     /clone_lib="Maize Mapping Project/DuPont Consensus
                     Library"
     note="this sequence is part of a project of EST
           assemblies resulting from the application of public
           contigs to seed DuPont contigs; this resource was
           assembled by DuPont as part of a collaboration for the
           overgo addressing of BACs in conjunction with the Maize
           Mapping Project"
BASE COUNT  655 a 442 c 478 g 621 t
ORIGIN

Query Match 55.9%; Score 1067.8; DB 11; Length 2196;
Best Local Similarity 76.5%; Pred. No. 9.7e-95;
Matches 1309; Conservative 0; Mismatches 402; Indels 0; Gaps 0;

Qy      13 GTTGTCCATGGAGAGAGGTGATAAATGACCCCTGTACTCCAAAACCTTAACCCCAAC 72
Db      134 GTTGCTCTCTGGGAAAAAGAGATCAATACTGTACTCTAGTACTCCAAAGCCCTTAATCCNAAT 193

Qy      73 CGGTTCACTTATGTGCGCTGAACCAAGTCAGAGCATGTTTCCAAACTGTTGATGCGGTT 132
Db      194 CGGTTCACTTATGAACATCAGCTAAACCGATCACCATTTCGAATGGTTGATGCGGTC 253

Qy      133 ATCCATGTTTATGCGGATAAGATTGTACGGAGAGCAATTAATCCTGTGGCTGATGCTACA 192
Db      254 ATTCATGTGTACCCCAATAAAGATTCTAAAGAAAGACTCTATCCAGTTGCCGATCAACT 313

Qy      193 ACCTTCTTCAGTCTGCATATATATCTCCGAGTAACGGCTGAGGGAACACAAGAAGACT 252
Db      314 ACCTTCTTTACTGATGATGCACTACGCTCTCTGTGTGTGGCTGCTGGGGATATCCGAACC 373

```

```

Qy      253 GTCTGCCATAATCGGTTAAATCTTCTTGAGCATAAAGTTTAAATTCATCTCTGATGTTAAAT 312
Db      374 GTGTGTCAATACCGGTTTAAATCTTCTAGAGCAAAAAATTAATCTTCACTTGATGTCAT 433

Qy      313 CGGATAGGGAATTTCTTGCCAGAGACTGCCCCACATCGTGATTTTAAATGTCAGG 372
Db      434 ACGGATAGAGAACTGCTTGTCTCAGAAAGCTGCACCTCACCGGAGCTTCTCAATGTTAGG 493

Qy      373 AAGGTTGACACTCATGTTTCATCTCAGCATGCAATCAAAAAACATCTGTGTAGGTTTC 432
Db      494 AAGGTTGATATCTCATGTTTCATCTCAGCATGCAATGAACAGAGAGCCTGTTGAGATT 553

Qy      433 ATAAATCCAACTAAAGAAAAAGAACCTGATCAGGTGGTTCATTTTCAGAGATGGTACTTAT 492
Db      554 ATCAAGTCTAAGTTGAGGAAGAACTGATCAGGTGTTGTAATCTTTAGAGATGGGACCTAT 613

Qy      493 ATGACTTTAAAGAGGTTTGTGAGAGCTTGACACTTAACCTGGGTATGATCTGAATGTTGAT 552
Db      614 TTGACTCTTAAGGAGGTTTGTGAGAGTTTGTGATTTGACAGGATATGACCTCAATGTTGAT 673

Qy      553 TTGCTAGATGCCATCAGACAAAAAGACATTTTCATCGTTTGTGACAAATTCATCTAAAA 612
Db      674 CTCTTAGATGTTCACTGATGAAGTACATTTTCATCGTTTGTGACAAAGTTCAATCTGAAA 733

Qy      613 TACAATCCATGTGGCCAAAGTAGGCTCAGAGAAATTTTCTCAACCAAGATAATCTTATT 672
Db      734 TATAATCCATGTGTCATCTCGCTGAGGAGATATTTTAAAGCAGGATAATCTTATT 793

Qy      673 CAAGCCGTTTTCTGCTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAA 732
Db      794 CAAGGTGCTTCTCTGCTGAGCTGACAAAAGAGATTTTCTGACCTTTGAAAGCAAGTAAA 853

Qy      733 TATCAGATGGCAGAAATATAGGATTTCAATCTACCGAAGGAAACAGAGTGAATGGAGCAA 792
Db      854 TATCAGATGGCTGAGTATCGGATATCTATTTATCGGAGGAAAGAAAGCAATGGAGCCAA 913

Qy      793 CTTCAGAGTTGGATAGTGAAACAATGCAAGTGGAAATGTTGTCTGGCTGGTTCCAG 852
Db      914 ATGGCAAGTTGGATAGTGAAACAATGAAATATATAGCAGAGATGTTGTTGGTTAAATCCAG 973

Qy      853 ATTCCAGCTTATATATATGTGTACAGAAATGGGTATCGTTACATCATCTTCAAAATCTT 912
Db      974 ATTCTCGGATATACAATGTATACAGGAGATGGGCAACAATCTCTTCCAAAACCTC 1033

Qy      913 CTTCAGCAACATTTCTGTTCTCTTTTGGAGTTACTATGATCCAGCTTCCACCCACAG 972
Db      1034 CTTCAGCAATATTTTCTCCCTCTTTTATGAAGTAACTGTTGACCCAGCTTTCATCTCCTCAG 1093

Qy      973 CTCCATGTTCTTCTGAAGCAGGTTGTAGGGTTGGACCTGTTGATGATGAAGTAAACCA 1032
Db      1094 CTCCATGTTTCTTAAACAGGTTGTTGGGTTGGATTTGGTGGATGATGAAGCAACCT 1153

Qy      1033 GAAAGGGCTCAACAAAGCAGCATGCCACACCTGAAACAGTGGACCAATGTGTTCAACCT 1092
Db      1154 GAAAGAGCCCAACAAACACATGCCAACACCTGAGCAATGAGCTAATATTTTCAATCCA 1213

Qy      1093 GCATTTTCATATTATGATGCTACTGCTATGCTAACTTATTCACCTCAACCAAGCTGGT 1152
Db      1214 GCCTATGCGTATATGATGCTATTTGCTATGCCAATCTGTATACACTGAACCAAGCTCGT 1273

Qy      1153 GAGTCAAGGGGAATGACCACTATCAAAATTCGTCACATGCTGGAGAGGCTGGAGATGTT 1212
Db      1274 GAGTCCAGGGCATGACAAACAATCTCGTCCACACTCTGGGGAGGCTGGTGATATC 1333

Qy      1213 GATCACTGGCAGCGACATTTCTCTCTGTCAACAATATGACATGGAATTAATCTAAGG 1272
Db      1334 GATCACCTTGCTGCGAGCATTTCTTACTTCAATAACATTTGCTCATGGGTGAACCTAAAG 1393

Qy      1273 AAGTCTCTGTGCTTTCAGTACTTGTACTATCTTGGTCAGATTGCTGTGGGATGTCGCCA 1332
Db      1394 AAATCCCTCTGTACTTACGTACCTGCTATTTATCTTCTGCTCAGATTGGTCTAGCAATGTCCTCC 1453

```

QY 1333 TTGAGCAACACCTCTTATTTCTTGACTATCATCGCAACCCCTTTTCCAAAGTCTTCTCAA 1392  
Db 1454 TTAAGTAACAACTCACTGTTTATCGACTATCATCGGAACCCCTTCCCAACATTTTCTTA 1513  
QY 1393 CGAGGTCTGAATGTCTCATATCTACGGATGACCCCTTCCAAATTCACCTGACAAAGAA 1452  
Db 1514 AGGGGTCTTAAGCTGTCTCTTTCGACTGACGCCCTCTACAAATTCACCTGACGAAGAA 1573  
QY 1453 CATTGGTGGAGAATACAGCATTTGCTGCTGCTGGAAGCTCAGTTCTTGATTTA 1512  
Db 1574 CTTTGGTGGAGAATACAGCTGTTGCTGCTTACATTGGAAGCTGAGCTCATGACTTG 1633  
QY 1513 TCGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTTTACATGCTCTCAAGGCGCAC 1572  
Db 1634 TGTGAATCGCTAGAAATCTGTCTACCAAGTCGGATTCTCTACAGGCTCAAGTCTCAC 1693  
QY 1573 TGAATTTGAAGAACTACTCTTCAAAGAGGACCTGCTGGAAATGATATTCACAGAACCAAT 1632  
Db 1694 TGGATCGGAGAACTACTACAGAAGGGTCCGGAGGCAATGACATCCACCAGACAAAT 1753  
QY 1633 GTACCCGACATCAGGGTTCAAATTTAGAGAGATGATCTGGAGAAATGAAATGAACTAGTG 1692  
Db 1754 GTTCTCCTACATCAGGGTGAATTTTCACACACTATCTGGAAGAAGAAATGGAGCTGATA 1813  
QY 1693 TACTCTGACAATGAGATCTTAATACCAGACG 1723  
Db 1814 CATTTTGGAAATGTTAACTACCAGGAAG 1844

RESULT 3  
CB642483  
LOCUS  
DEFINITION  
OSJNEB02K24.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Jantasiyirarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,B., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang G.  
Large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea  
Unpublished  
Contact: Rod Wing  
Arizona Genomics Institute  
University of Arizona  
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ 85721-0088, USA  
Tel: 520 626 3967  
Fax: 520 621 9288  
Email: http://genome.arizona.edu  
PCR Primers  
FORWARD: gta aac cga cgg cca gtc  
BACKWARD: gga aac agc tat gac cat g  
Plate: 02 row: K column: 24  
Seq primer: gta aac cga cgg cca gtc  
Location/Qualifiers  
1..832  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="OSJNEB02K24"  
/tissue\_type="Leaf"  
/dev\_stage="3 week"  
/lab\_host="DH10B"  
/clone\_lib="OSJNEB"  
/note="Vector: pBluescript II KS +; Site\_1: EcoRI; Site\_2: EST."

BASE COUNT 245 a 169 c 169 g 249 t  
ORIGIN  
Query Match 35.7%; Score 682.8; DB 14; Length 832;  
Best Local Similarity 88.9%; Pred. No. 2.1e-57;  
Matches 738; Conservative 0; Mismatches 92; Indels 0; Gaps 0;  
QY 157 TGTACGGAGAGCATTTATCTGCTGCTGATGCTCAACCTCTTCTCAGCTGACTTGCATTAT 216  
Db 2 TATACGGAAATACCTTTATCTGCTGCTGATGCGACACCTCTTTTACTGACTTGCATTAT 61  
QY 217 ATTCTCCAGTAACCGGCTGCGAGGAACAACAAGAACTGTCTGCCATAATTCGGTTAAATCTT 276  
Db 62 ATGCTCCGTGTGATTTCTGCTGGAAACAACAAGAACCGTGTGCCATAACCGGTTAAATCTT 121  
QY 277 CTTGAGCATATAATTTAAATTCATCTGATTTAAATTCGGATAGGGAATTTCTTCCCCAG 336  
Db 122 TTAGAGCATATAATTTCAAATTTTCAATTTGATTTAAACCGGATAGGAGTTTCTTCCCCAG 181  
QY 337 AAGACTGCCACATCTGATTTTACAAATGTCAGGAAGGTTGACATCATGTTTCATCAT 396  
Db 182 AAGACTGCCACATCTGATTTTACAAATGTCAGGAAGGTTGATCTCATGTACACCAT 241  
QY 397 TCAGCATGATGAATCAAAAACATCTGTTGAGGTTTCATAAAATCCAAAATCAAGAAAGAA 456  
Db 242 TCAGGCTGATGATCAAAAACATCTGCTGAGATTTATAAATCCAAATCAAGAAAGAA 301  
QY 457 CTTGATAGGTGGTTCATTTTTCAGAGATGGTACTTATATGACTTTTAAAGAGAGGTTTTGAG 516  
Db 302 CTTGATAGGTGGTTCATTTTTCAGAGACGGCACTTATATGACTCTCAAGAGAGGTTTTGAG 361  
QY 517 AGCTTGGACTTAACTGGGATGATCTGAATGTTGATTTGCTAGATGTCATCAGACACAA 576  
Db 362 AGCTTGGACTTAACTGGGATGATCTGAATGTTGATTTTACTAGATGTCATCAGGACAA 421  
QY 577 AGCATTCTCATCGTTTTCAGAAATTCATCTTAAATCAATTCATTCAGTGGCCAAAGTAGG 636  
Db 422 AGTACATTTTCATCGTTTTCAGCAATTCACCTTAAGTACAATCCGTGTCGCAAGTAGG 481  
QY 637 CTCAGAGAAATTTTCTCAAAAGATTAATCTTATTCAGGCGGTTTTCTTCTGCTGAGTTG 696  
Db 482 CTCAGAGATTTTCTCAAAAGATTAATCTTATTCAGGCGGTTTTCTTCTGCTGAGTTG 541  
QY 697 ACAAGCAAGTTTCTCTGACCTTCTGCTAGCAATATCAGATGCGAGATATAGGATT 756  
Db 542 ACAAGCAAGTTTCTCTGACCTTCTGACCAAGTAAATACCAGATGGCTGAATATAGGATT 601  
QY 757 TCAATCTACGGAAGGAACAGAGTGAATGGGACCAACTTGCAGTGGATAGTGAACAAAT 816  
Db 602 TCAATCTATGGGAGGAAGCAAGTGAATGGGACCACTTGCAGCTGGATAGTGAACAAAT 661  
QY 817 GAATTTGCAAGTGGAAATTTGCTGCTGGCTGTTGAGATTCACGCTTATATATATGTTGATC 876  
Db 662 GAATTTGCAAGTGGAAATTTGCTGCTGGCTGTTGAGATTCACGCTTATATATATGTTGATC 721  
QY 877 AAGGAATGGGTATCGTTACATCATCTTCCAAATCTTCTTGCAACAAATTTTCTTCTCTT 936  
Db 722 AAGGAATGGGTATTTGTTACATCATCTTCCAAATCTTCTTGCAACAAATTTTCTTCTCTT 781  
QY 937 TTTGAGGTACTATTGATCCAGCTTTCACACCCACAGCTTCATCTCTCTCTCTCTCTCT 986  
Db 782 TTTGAGGTACTATTGATCCAGCTTTCATCTCCGAGCTTCAGCTCTCTCTCTCTCTCT 831  
RESULT 4  
CB643461  
LOCUS  
DEFINITION  
OSJNEB04C17.f OSJNEB Oryza sativa (japonica cultivar-group) cDNA  
clone OSJNEB04C17 5', mRNA sequence.  
ACCESSION  
VERSION  
CB643461  
GI:29638452

```

KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)

REFERENCE 1 (bases 1 to 798)
AUTHORS Jantasuriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E.,
Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.
TITLE Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
JOURNAL Unpublished
COMMENT Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtg
BACKWARD: gga aac agc tat gac cat g
Plate: 04 row: C column: 17
Seq primer: gta aaa cga cgg cca gtg.

FEATURES
source
1..798
Location/Qualifiers
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="OSUNEB04C17"
/tissue_type="Leaf"
/dev_stage="3 week"
/lab_host="DH10B"
/clone_lib="OSUNEB"
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:
XhoI; 24 hrs after inoculation with Rice Blast (Che 86061
)"

BASE COUNT 242 a 154 c 166 g 236 t
ORIGIN
Query Match 34.1%; Score 651.4; DB 14; Length 798;
Best Local Similarity 88.6%; Pred. No. 2.3e-54;
Matches 706; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 157 TGTACGGAGCATTTATCTCTGGCTGATGCTACAACTCTCTCAGTGGTTCATTAT 216
DB 2 TATACGGAATACCTTATTAATGTTCTGATGCGAACCTCTTACTGACTTGCATTAT 61

QY 217 ATTCTCCGAGTAACGGCTGCGAGGAACACAGAACTGTCTGCATAATCGGTTAAATCTT 276
DB 62 ATGCTCCGGTGATTTCTGCTGGAACACAGAACCGGTGCGATACCGGTTAAATCTT 121

QY 277 CTTGACATAGTTTAAATTCATCTGATGTTAAATGCGGATAGGAATTTCTGCCGAG 336
DB 122 TTAGAGCATAAATTCAAATTTTCATTTGATGTTAAACGCGGATAGGAGTTCTTCCCGAG 181

QY 337 AAGACTGCCCCACATCGTGATTTTACATGTCAGGAAGTTTGACATCATGTTTCATCAT 396
DB 182 AAGACTGCACCATCGTGATTTTACATGTCAGGAAGTTTGATCATGATACACCATC 241

QY 397 TCAGATGCATGAATCAAAAACATCTGTTGAGGTTTCATAAAATCAAACTAAGAAAAGAA 456
DB 242 TCAGCGTGATGATCAATCAAAAACATCTGCGATTTATAAATCCAACTAAGGAAGAA 301

QY 457 CTGATGAGGTGGTCATTTTCAGAGAGTGATCTTATATGACCTTTAAAGAGGTTTTGAG 516
DB 302 CCTGATGAGGTGTCATTTTCAGAGACGCGCACTTATATGACCTCTCAAGAGGTTTTGAG 361

QY 517 ACCTTGGACTTAACCTGGGTATGATCTGAATGTTGATTTGCTAGATGTCATGACACAAA 576
DB 362 AGCTTGGACTTAACCTGGGTATGATCTGAATGTTGATTTTACTAGATGTCATGCGGACAAA 421

QY 577 AGCATTTCATCGTTTTCACAAAATTCATCTTAAATAATCAATCCATGTCGCAAAAGTAGG 636
DB 422 AGTACATTTCATCGTTTTCACAAAATTCACCTTAAGTACAATCCGTGTGGCCAAAGTAGG 481

QY 637 CTCAGAGAAATTTTCCTCACAACAGATATCTTATTAAGCCGCTTTCTTGTCTGAGTTG 696
DB 482 CTCAGAGAGATTTTCCTCAAAACAGAGCAATCTTATCAAGGTCGTTTCTTGTCCGAGCTA 541

QY 697 ACAAGCAAGTTTCTCTGACCTTCTCTAGCAAAATATCAGATGGCAGAAATATAGGATT 756
DB 542 ACAAGCAAGTTTCTCTGACCTTACACAGTAGTAATACAGATGGCTGAATATAGGATT 601

QY 757 TCAATCTACGGAAGGAAACAGAGTGAATGGACCAACTTTCGAAGTTGATAGTGAACAAT 816
DB 602 TCAATCTATGAGGAGGAAGCAAAAGTGAATGGACCAACTTTCGAAGTGGATAGTGAACAAT 661

QY 817 GAATTCACAGTGGAAATGTTCTGCGCTGCTTCAGATTCACCGCTTATATATGTTGATC 876
DB 662 GAATTCACAGTGGAAATGTTCTGCGCTGCTTCAGATTCACCGCTTATATATGTTGATC 721

QY 877 AAGAAATGGGTATGCTTACATCATCTTCAAAATCTTCTTGACAACTTTTCGTTCTCTT 936
DB 722 AAGAAATGGGTATGTTTACATCATCTTCAAACTCTTCTTGACAACTTTTCGTTCTCTT 781

QY 937 TTTGAGGTTACTATGA 953
DB 782 GTTGAGGTCACCATTTGA 798

RESULT 5
AW017903 624 bp mRNA linear EST 13-SEP-1999
LOCUS 614020B05.y1 614 - root cDNA library from Walbot Lab Zea mays cDNA,
DEFINITION mRNA sequence.
ACCESSION AW017903
VERSION AW017903.1 GI:5871432
KEYWORDS EST.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 624)
AUTHORS Walbot,V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL Unpublished
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614020 row: B column: 05.
FEATURES
Location/Qualifiers
1..624
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="W23"
/db_xref="taxon:4577"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XJOLR"
/clone_lib="614 - root cDNA library from Walbot Lab"
/note="Organ: root; Vector: pBluescriptII SK+; Site 1:
EcoRI; Site 2: XhoI; 3-4 days old root tissue from Walbot
Lab (LM)"

BASE COUNT 164 a 149 c 133 g 176 t 2 others
ORIGIN
Query Match 31.7%; Score 606; DB 9; Length 624;

```



```
Db 664 GGTCTTGCCATGCTCCCTTGAGCAACAACTCATTTGATTAATATCACCGAACCCCT 723
QY 1375 TTTCACAGCTTCTTCAACGAGGCTGTAATGCTCATTAATACGAGTACCCCTTTGCAA 1434
Db 724 TTCCCAACATTTTCTCTAAGAGCCCTTAACGTTTCTCTATCAACCGATGACCCCTTTGCAA 783
QY 1435 ATTACCTGACAAAAGAACCACTTGTGGAGAAATACAGCATTCGCTTCGCTGTGGAAG 1494
Db 784 ATTACCTGACAAAAGAACCACTTGTGGAGAAATATAGCATCGCTTCGCTGTGGAAG 843
QY 1495 CTCAGTCTCTTGATTTATGCGAAATTTGCGAGGA 1528
Db 844 CTAAGTTCATGCGACCTATGTGAATTTGCTAGGA 877

RESULT 7
LOCUS AW671750 634 bp mRNA linear EST 19-JUL-2000
DEFINITION LGL_351_F02.bi_A002 Light Grown 1 (LGI) Sorghum bicolor cDNA, mRNA
ACCESSION AW671750
VERSION AW671750.1 GI:7535646
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 634)
Cordonnier-Pratt M.-M., Gingle A., Marsala, C. and Pratt, L.H.
An EST database from Sorghum: light-grown seedlings
Unpublished
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 516
POLYA=No.

FEATURES
source
1..634
/organism="Sorghum bicolor"
/mol_type="mRNA"
/db_xref="taxon:4558"
/clone_lib="Light Grown 1 (LGI)"
/note="Organ: 10- to 14-day-old light-grown (greenhouse)
seedlings; Vector: Lambda Zap; Site 1: XhoI; Site 2: EcoRI
; The library was made from poly-A RNA in the cloning
vector lambda Zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT 190 a 137 c 126 g 181 t
ORIGIN

Query Match 30.2%; Score 576.8; DB 9; Length 634;
Best Local Similarity 96.0%; Pred. No. 4.1e-47;
Matches 603; Conservative 0; Mismatches 22; Indels 3; Gaps 1;

QY 523 GACTTAACCTGGGTATGATCTGAATGTTGATTTGCTAGATGTCATCGCAGACAAAAGCACA 582
Db 7 GCCTTAACCTGGGTATGATCTGAATGTTGATTTGCTAGATGTCATCGCAGACAAAAGCACA 66
QY 583 TTTCATCGTTTGGACAAATTCATCTAAATAATCAATCCATGTGGCCAAAGTAGGCTCAGA 642
Db 67 TTTCATCGTTTGGACAAATTCATCTAAATAATCAATCCATGTGGCCAAAGTAGGCTCAGA 126
QY 643 GAAATTTTCCTCAACACAGATTAATCTTATTCAGGCCGTTTTCTTGCTGAGTTGACAAAG 702
```

```
Db 127 GAAATTTTCCTCAACACAGCAATCTCATTCAGGCCGTTTTCTTGCGAGTTGACAAAG 186
QY 703 CAAGTTTTCTCTGACCTTTTCTGTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATC 762
Db 187 CAAGTTTTCTCTGACCTTTTCTGTAGCAAAATATCAGATGGCAGAAATATAGGATTTCAATC 246
QY 763 TACGAGGAGAAACAGAGTGAATGGACCAACTTCGAAGTTGGATAGTGAACAATGAATTG 822
Db 247 TATGGGAGGAAACAAAGTGAATGGACCAACTTCGAAGTTGGATAGTGAACAATGAATTG 306
QY 823 CACAGTGGAAATGTTGCTGCTGCTTCCAGATTCACGCTTATATATATGTTGACAAAGAA 882
Db 307 CACATGAAATGTTGCTGCTGCTTCCAGATTCACGCTTATATAGTGTACAGGGA 366
QY 883 ATGGGTATCGTTACATCATTCCAAATCTTCTTGACAAACATTTTCCTCTTTTTCAG 942
Db 367 ATGGGTATGTTGCTGCTGCTTCCAAATCTTCTTGACAAACATTTTCCTCTTTTTCAG 426
QY 943 GTTACTATTGATCCAGCTTCACACCCACAGCTCCATGCTTCTTGAGGACGAGTTGTAGGG 1002
Db 427 GTTACTATTGATCCAGCTTCACACCCACAGCTCCATGCTTCTTGAGGACGAGTTGTAGGG 486
QY 1003 TTGGACCTGGTTGATGATGAAAGTAAACCAAGAGGCGTCCAAACAAAG--CACATGCC 1059
Db 487 CTGACCTGGTTGATGATGAAAGTAAACCAAGAGGCGTCCAAACCAACCATGCT 546
QY 1060 ACACCTGAACAGTGGACCAATGTTTCAACCTGCATTTTCATATTATGCTACTACTGC 1119
Db 547 ACACCTGAACAGTGGACCAATGTTTCAACCTGCATTTTCATATTATGCTACTACTGC 606
QY 1120 TATGCTAACTTATTCACCTTAAACAGC 1147
Db 607 TATGCTAACTTATTCACCTTAAACAGC 634

RESULT 8
LOCUS CB668055 845 bp mRNA linear EST 09-APR-2003
DEFINITION OSJNEd15108.f OSJNEd Oryza sativa (japonica cultivar-group) cDNA
clone OSJNEd15108 5', mRNA sequence.
ACCESSION CB668055
VERSION CB668055.1 GI:29671780
KEYWORDS EST.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretidoideae; Oryzaceae; Oryza.
1 (bases 1 to 845)
Jantasuriyarat, C., Lu, G., Gowda, M., Hatfield, J., Zhou, B., Mazur, E.,
Kudrna, D., Dean, R., Soderlund, C., Wing, R. and Wang, G.
Large-scale identification of ESTs involved in the interaction
between rice and Magnaporthe grisea
Unpublished
Contact: Rod Wing
Arizona Genomics Institute
University of Arizona
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ
85721-0088, USA
Tel: 520 626 3967
Fax: 520 621 9288
Email: http://genome.arizona.edu
PCR Primers
FORWARD: gta aaa cga cgg cca gtc
BACKWARD: gga aac agc tat gac cat g
Plate: 15 row: L column: 08
Seq primer: gta aaa cga cgg cca gtc.
Location/Qualifiers
1..845
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
```





```

QY 312 TCGGGATAGGAATTTCTTGGCCAGAGAGACTGCCCCACATCGTGATTTTACAATGTCAG 371
Db 310 -----GGAGTTTCTTGGCCAGAGAGACTGCCCCACATCGTGATTTTACAATGTCAG 360
QY 372 GAAGGTTGACACTCATGTTTCATCATTGACATGATGAATCAAAACATCTGTGAGGTT 431
Db 361 GAAGGTTGATATCATGTATACACCTACGCGTGATGAATCAAAACATCTGTGAGATT 420
QY 432 CATAAATCCAACTAAGAAAAGAACCTGATGAGGTGGTCAITTTTCAGAGATGGTACTTA 491
Db 421 TATAAATCCAACTAAGGAAAGAACCTGATGAGGTGGTCAITTTTCAGAGAGCGCACITTA 480
QY 492 TATGACTTTAAAGAGGTTTTTGAGAGCTTGGACTTAACCTGGGTATGATCTGAATGTTGA 551
Db 481 TATGACTCTGAAGAGGTTTTTTGAGAGCTTGGACTTAACCTGGGTATGACTTGAATGTTGA 540
QY 552 TTTGCTAGATGTCATGACAGAAAAGACATTTTCATCGTTTTCAGAAAATTCAACTCAA 611
Db 541 TTTTACTAGATGTCATGCGGCAAAAGTACATTTTCATCGTTTTCAGAAAATTCAACTCAA 600
QY 612 ATACATCCATGTCGGCCAAAGTAGGCTCAGAGAAATTTTCTCCTCAAAAGAGATAATCTTAT 671
Db 601 GTACATCGTGTGGCCAAAGTAGGCTCAGAGAGATTTTCTCCTCAAAAGAGATAATCTTAT 660
QY 672 TCAAGCCGTTTTTCTGTGAGTTGACAAAGCAAGTTTCTCTGACCTTTCTGTGAGCAA 731
Db 661 CCAAGGTCGTTTTTCTGTGCGAGCTAACAAAGCAAGTTTCTCTGACCTTTACAGCAAGTAA 720
QY 732 ATATCAGATGGCAGATATAGGATTTCAATCTACGG 767
Db 721 ATACCAGATGGCTGAATATAGGATTTCAATCTATGG 756

```

## RESULT 10

```

BM359276
LOCUS BM359276
DEFINITION GA_Ea0018J03r Gossypium arboreum 7-10 dpa fiber library Gossypium
          arboreum cDNA clone GA_Ea0018J03r, mRNA sequence.
ACCESSION BM359276
VERSION   BM359276.1
KEYWORDS  EST.
SOURCE    Gossypium arboreum
          Gossypium arboreum
          Gossypium arboreum

```

## REFERENCE

```

AUTHORS  Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
          ,D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE    An integrated analysis of the genetics, development, and evolution
          of the cotton fiber
JOURNAL  Unpublished
COMMENT  Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Total High Quality bases = 590
          Seq primer: TAATACGACTCACTATAGG
          High quality sequence stop: 728.

```

## FEATURES

```

source
1..872
/organism="Gossypium arboreum"
/mol_type="mRNA"
/strain="AKA"
/cultivar="8400"
/db_xref="taxon:29729"
/clone="GA_Ea0018J03r"
/tissue_type="Fibers isolated from bolls harvested 7-10
          dpa"

```

RESULT 11  
BU973129

```

/lab_host="E. coli"
/clone_lib="Gossypium arboreum 7-10 dpa fiber library"
/notice=vector: pbk-CMV; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 257 a 174 c 164 g 271 t 6 others
ORIGIN
Query Match 27.1%; Score 517; DB 12; Length 872;
Best Local Similarity 75.4%; Pred. No. 2e-41;
Matches 651; Conservative 0; Mismatches 211; Indels 1; Gaps 1;

```

```

QY 574 AAAAGCAGCATTTTCATCGTTTGGACAAATTCATCTAAATAACAATCCATGCGCCAAAGT 633
Db 2 AAGAGCAGCTTTTCATCGATTGGACAAATTCATCTAAATAACAATTCCTTTGGGCAAGC 61
QY 634 AGGCTCAGAGAAATTTTCTCAAAACAAGATATCTTATTCAGAGCGGTTTTCTTGCTGAG 693
Db 62 AGACTGAGAGAGATCTTTTAAAGCAGGACAACTCTTATCCAAGGAGGTTTTCTGCGAGAA 121
QY 694 TTGACAAAGCAAGTTTTCTCTGACCTTCTGCTAGCAAAATATCAGATGGCAGAAATATAGG 753
Db 122 GTTACAAAGCAAGTGCTATCAGATCTTGAAACAAGCAAAATATCAGATGGCAGAGTACAGG 181
QY 754 ATTTCAATCTACGGAAGGAAACAGAGTGAATGGGACCAACTTTCGAAGTTGATAGTGAAC 813
Db 182 GTGTCCATATATGAGAGAAACAAGTGAATGGGACCAAGTTGGCAAGTTGGTTATTATAC 241
QY 814 AATGAATTCACAGTGGAAATTTTGTCTGGCTGGTTCCAGATTCCACGCTTATAAATGTG 873
Db 242 AATGAATTTATAGTGAGACTTACTGTTTGGTTAATCCAGTTTACCAGATTGTATAATGTG 301
QY 874 TACAAGAAATGGGTATCGTTTACATCATTCGAAATCTTCTTGACAAATTTCTGTCCTT 933
Db 302 TACAAGCAATGGGAATTTGTAATCTTTTTCAGAAATATCTTGGACAAATGTTTTCATTCCA 361
QY 934 CTTTGTGAGTTTACTATTGATCCAGCTTCACCCAGAGCTCCATGCTTCTCTGAGAGCAG 993
Db 362 CTATTGAGTTACAGTTGATCCAAATTTCTCATCTCAACTACACGTTGTTCTCTAAAATG 421
QY 994 GTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACAGAAAGCGTCCCAACAAAGCAC 1053
Db 422 GTGGTTGGCTTTGATTAGTTGATGAGAGTAAACAGAGAGCGCTCTTACAAGCAC 481
QY 1054 ATGCCACACCTGGAACAGTGGACCAATGTGTTCAACCTCGCATTTTCATATTATGCGTAC 1113
Db 482 ATGCCAACTCTGCTGAATGAGTAAATGAAATTTAATCTGCTATTTCTTATTATGCTTAC 541
QY 1114 TACTGCTATGCTAACTTATTTCACCTTAAACAGCTGGTGAATCAAGGGGAATGACCACT 1173
Db 542 TACTTTTATGCAAACTTTTATACACTCAACAAGCTTCGTGAGTCAAAAGGAATGCAGACA 601
QY 1174 ATCAAAATTCGTCACATGCTGGAGAGCTGGAGATGTTGATCACTTGGCAGCGACATTT 1233
Db 602 ATAAACTCCGACCTCACTGTGGAGGCTGGTATATTGACCATNTAGCTGCTGCTTTC 661
QY 1234 CTTCTGTGTCAACAATATCATGGAATTAATCTAAGGAAGTCTCTGTGCTTTCAGTAC 1293
Db 662 CTTCTGTG-CACANTATATCTCACGGGATTAATCTNCGAAATCCCTGTTTTCAGTAC 720
QY 1294 TTGTACTATCTTGTGTCAGATTGGTCTCGCGATGTCCCATTTGACCAACAACCTCTTATT 1353
Db 721 TTGTATTACCTTCGCTCAGATCGGATNGGCAATGTCCTCTGAGCAATAATATCCCTTTTC 780
QY 1354 CTTGACTATCATCGCAACCTTTTCCAAACGTTCTTCCAAAGAGTCTGAAATGCTCATTA 1413
Db 781 CTGACTATCATCGCAACCTTTTCTTCACTTTTCAGGGGCGCTAAATGTTTCTCTT 840
QY 1414 TCTACGAGTACCTTTTGCAAAT 1436
Db 841 TCATCTGATGACCTTAAACAAAT 863

```

LOCUS BU973129 647 bp mRNA linear EST 22-OCT-2002  
 DEFINITION HB23N17: BC Hordeum vulgare subsp. vulgare cDNA clone HB23N17  
 5-PRIME, mRNA sequence.  
 ACCESSION BU973129  
 VERSION BU973129.1 GI:24223322  
 SOURCE EST.  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 647)  
 REFERENCE Radchuk,V., Zhang,H., Weschke,W., Potokina,E. and Wobus,U.  
 AUTHORS Barley ESTs from developing seeds  
 TITLE Unpublished  
 JOURNAL  
 COMMENT Contact: Stein Nils  
 Molecular Markers Group, Department Genbank  
 Institute of Plant Genetics and Crop Plant Research (IPK)  
 Corrensstr. 3, 06466, Gatersleben, Germany  
 Tel: 039482-5522  
 Fax: 039482-5595  
 Email: stein@ipk-gatersleben.de  
 Insert Length: 647 Std Error: 0.00  
 Plate: 23 row: N column: 17  
 Seq primer: M13rev.

## FEATURES

Location/Qualifiers  
 1..647  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="barke"  
 /db\_xref="GABI:237971"  
 /db\_xref="taxon:112509"  
 /clone="HB23N17"  
 /tissue\_type="developing caryopsis"  
 /dev\_stage="8-15 DAP (days after pollination)"  
 /lab\_host="XL10-Gold"  
 /clone\_lib="BC"  
 /notes="Vector: pBluescript SK+; Site 1: EcoRI (5'-end of cDNA); Site 2: XhoI (3'-end of cDNA); developing caryopsis  
 , 8-15 DAP(days after pollination) Due to a cloning  
 artefact caused by the kit, in most cases the EcoRI site  
 is NOT present, as well as the EcoRI adapter used for  
 cloning. To excise the insert, restriction sites upstream  
 EcoRI should be used (e.g. BamHI, SalI, PstI). NOTE: Also  
 due to the cloning system used Blue/white selection for  
 recombinants is not 100% reliable."

## BASE COUNT

176 a 148 c 136 g 187 t

## ORIGIN

Query Match 26.6%; Score 508.2; DB 13; Length 647;  
 Best Local Similarity 87.1%; Pred. No. 1.7e-40;  
 Matches 558; Conservative 0; Mismatches 83; Indels 0; Gaps 0;  
 QY 1073 GGACCAATGTGTTCAACCCCTGCATTTTCATATTATGCGTACTACTGCTAACTTAT 1132  
 Db 1 GGACCAATGTGTTCAACCCCTGCATTTTCATATTATGCGTACTACTGCTAACTTAT 60  
 QY 1133 TCACCCTAACACAGTCGCTGAGTCAAGGAATGACCACTATCAAAATTCGCTCCACATG 1192  
 Db 61 ACACACTGAACAAGCTGCGTGAAGTCAAGGGGATGAATATCAAAATTCGCTCCACATG 120  
 QY 1193 CTGAGAGGCTGGAGATGTTGATCCTTGGCAGCAGCATTTCTTCTGTGCACAAATAT 1252  
 Db 121 CTGTGAGGCTGGAGATGTTGATCCTTGGCAGCAGCATTTCTTCTGTGCACAGTATAT 180  
 QY 1253 CACATGGAATTAATCTAAGGAAGTCTCCTGTGCTTCAAGTACTTGTACTATCTTGTGTCAGA 1312  
 Db 181 CACATGGAATCAATTTAAGGAAGTCTCCTGTGCTTCAAGTACTTGTACTATCTTGTGTCAGA 240  
 QY 1313 TTGGTCTGGCGATGTCGCCATTGAGCAACCACTCTTATTTCTTGTACTATCATCGCAACC 1372  
 Db 241 TCGGTCTAGCAATGTGCCCAATTAAGCAACAACCTCTTGTCTTGTGATACCATCGGAACC 300

QY 1373 CTTTTCACACGTTCTTCCACAGAGTCTGAATGCTCAATATCTACGATGACCCCTTGC 1432  
 Db CTTTTCCTATGTTTTTCCACAGAGACTGAATGCTCAATGCTCCACGATGATCCATGTC 360  
 QY 1433 AAATTCACCTGACAAAAGAACCATTTGGTGAAGATACAGCATTCCTCTTCGCTGTGGA 1492  
 Db AAATTCATCTGACAAAAGAGCCGTTGGTGCAGGAATACAGCATTCGCCCTTCGCTGTGGA 420  
 QY 1493 AGCTCAGTCTCTTGATTTATGCGAAATTCGAGGAACCTCTGTTTACCAATCTGGGTTT 1552  
 Db AGCTCAGTCTCTTGATTTATGCGAAATTCGAGGAATTCGTTGATCAATCAGGCTTTT 480  
 QY 1553 CACATGCTCTCAAGCGGCACCTGGATTGTAAGAACTACTTCAAAAGAGAGACTGCTGAA 1612  
 Db CACATGCTCTCAAGCGGCACATTTGGATTGGCAAGAACTACTTCAAGAGAGGCCCTCAGGA 540  
 QY 1613 ATGATATTTCAGAAACCAATGTACCGCACATCAGGTTCAATTTAGAGAGATGATCTGA 1672  
 Db ACGATATATCCAGAACGAATGTGCCACCACATCAGGATCGAATTCAGGACCTGATCTGA 600  
 QY 1673 GAAATGAATGAAACTAGTCTACTCTGACAATGAGATCTTA 1713  
 Db GAGAGAAATGCGAGCTCGTTTACCTTGATACGTCATCTTA 641

## RESULT 12

BI955474  
 LOCUS HVSMEM0023E15f 676 bp mRNA linear EST 19-OCT-2001  
 DEFINITION HVSMEM0023E15f Hordeum vulgare green seedling EST library  
 HVCNDA0014 (Blumeria infected) Hordeum vulgare subsp. vulgare cDNA  
 clone HVSMEM0023E15f, mRNA sequence.  
 ACCESSION BI955474  
 VERSION BI955474.1 GI:16301827  
 KEYWORDS EST.  
 ORGANISM Hordeum vulgare subsp. vulgare  
 Hordeum vulgare subsp. vulgare  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae  
 ; Triticeae; Hordeum.  
 1 (bases 1 to 676)  
 REFERENCE Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D.,  
 Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons  
 ,J., Oates,R. and Main,D.  
 TITLE Development of a genetically and physically anchored EST resource  
 for barley genomics: Blumeria infected Morex (compatible) seedling  
 cDNA library  
 JOURNAL Unpublished  
 COMMENT Contact: Wing RA  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Tel: 864 656 7288  
 Fax: 864 656 4293  
 Email: rwing@clemson.edu  
 Total hg bases = 442  
 Seq primer: AATTAACTCTACTAAAGGG  
 High quality sequence start: 23  
 High quality sequence stop: 541.  
 Location/Qualifiers  
 1..676  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Morex"  
 /db\_xref="taxon:112509"  
 /clone="HVSMEM0023E15f"  
 /tissue\_type="green seedling leaf"  
 /lab\_host="TJCI21"  
 /clone\_lib="Hordeum vulgare green seedling EST library  
 HVCNDA0014 (Blumeria infected)"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Morex (mla) plants were greenhouse grown in the R  
 Wise lab at Iowa State University, Ames, IA; 7 day old  
 green seedlings were infected with isolate 5874 of

## FEATURES

source  
 1..676  
 /organism="Hordeum vulgare subsp. vulgare"  
 /mol\_type="mRNA"  
 /cultivar="Morex"  
 /db\_xref="taxon:112509"  
 /clone="HVSMEM0023E15f"  
 /tissue\_type="green seedling leaf"  
 /lab\_host="TJCI21"  
 /clone\_lib="Hordeum vulgare green seedling EST library  
 HVCNDA0014 (Blumeria infected)"  
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:  
 XhoI; Morex (mla) plants were greenhouse grown in the R  
 Wise lab at Iowa State University, Ames, IA; 7 day old  
 green seedlings were infected with isolate 5874 of

*Blumeria graminis* f. sp. *hordei*, and leaves were harvested 24, 48 and 72 hr post-inoculation and snap frozen (Wise). In the TU Close lab at the University of California, Riverside, total RNA was prepared from each sample pool, equal quantities of all three RNA pools were combined, poly(A) RNA was purified from the mixture, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give phagescript SK(-) cDNA phagemids (Chin). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TU, Wing R, Kleinhoof A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. *Barley Genetics Newsletter* 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)"

```

BASE COUNT      186 a   132 c   153 g   205 t
ORIGIN

Query Match      26.6%; Score 507.6; DB 12; Length 676;
Best Local Similarity 87.6%; Pred. No. 1.9e-40;
Matches 578; Conservative 0; Mismatches 79; Indels 3; Gaps 2;

QY 777 GAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 836
Db 18 GAGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 77

QY 837 TGCTGGCTGGTTTCAGATTCACAGCTTATATATATGTGTACAAAGAAATGGGTATCGTTAC 896
Db 78 TGTTGGCTGATTCAGATTCACAGCTTATATATACGTGTACAAAGAAATGGGTATCGTTAC 137

QY 897 ATCATTCCAAAATCTTTTGACAACAATTTTCGTTCTCTTTTGGAGTTACTATTGATCC 956
Db 138 ATCATTCCAAAATCTTTTGACAACAATTTTCGTTCTCTTTTGGAGTTACTATTGATCC 197

QY 957 AGCTTCACACCAACAGCTCCATGTTCTCTGAGCAGGTGTAGGGTTGGACCTGGTTGA 1016
Db 198 AGCTTCGACCCGCGAGCTTCATGTTCTCTCTAAAGCAGGTGTAGGGTTGGATTTGGTTGA 257

QY 1017 TGATGAAGTAAACAGAAAGCGCTCCAAACAAAGCACATGCCACACCTGACAGTGGAC 1076
Db 258 TGATGAGATTAACCTGAAAGCGTCCAACTAAGCATATGCTCACCCTGAAAGATGGAC 317

QY 1077 CAATGTGTTCAACCCCTGCAATTTTCATATATATGCTACTACTGCTATGCTAACTTATTCAC 1136
Db 318 AAATGTCTTCAACCCCTGCAATTTTCATATATATGCTACTACTGCTATGCTAACTTATTCAC 377

QY 1137 CCTAAACAGCTGGTGAAGTAAAGGAATGACCATATCAAAATCCGTTCCACATGCTGG 1196
Db 378 ACTGAACAGCTGGTGAAGTAAAGGAATGACCATATCAAAATCCGTTCCACATGCTGG 437

QY 1197 AGAGGCTGAGATGTTGATCACTTCGACGACATTTCTCTCTCTGTCACACATATACCA 1256
Db 438 TGAGGCTGAGATGTTGATCACTTCGACGACATTTCTCTCTCTGTCACAGTATATACCA 497

QY 1257 TGAATTAATCTAAGGAAGTCTCTGTGCTTCAGTACTTGTACTATCTTGGTCAGATTGG 1316
Db 498 TGAATTAATCTAAGGAAGTCTCTGTGCTTCAGTACTTGTACTATCTTGGTCAGATTGG 557

QY 1317 TCTGGGATGTCCTTATGAGCAACAATCTCTTTATTTCTTGACTATATCGCAACCCCTTT 1376
Db 558 GGTAGCAATGTCCTTATGAGCAACAATCTCTTTATTTCTTGACTATATCGCAACCCCTTT 617

QY 1377 TCCAACGTTCTTCCAACGAGTCTCAATGCTCATATCTAGGATGACCCCTTTCGAAT 1436
Db 618 CCTA--TGTTTTTCAACGAGGACTGAATG-GTTATTTGCCAGGATGATCCATTGCAAT 674

```

# RESULT 13 BF268586

```

LOCUS              880 bp      mRNA      linear      EST 07-MAR-2001
DEFINITION        GA_EB0002D19f Gossypium arboreum 7-10 dpa fiber library Gossypium
                    arboreum cDNA clone GA_EB0002D19f, mRNA sequence.
ACCESSION        BF268586
VERSION          BF268586.2 GI:13247099
KEYWORDS          EST.
SOURCE            Gossypium arboreum
ORGANISM          Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
REFERENCE        1 (bases 1 to 880)
AUTHORS          Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
                  D., Wood,T.C., Leslie,A. and Wilkins,T.A.
TITLE            An integrated analysis of the genetics, development, and evolution
                  of the cotton fiber
JOURNAL          Unpublished
COMMENT          On Nov 17, 2000 this sequence version replaced gi:11199581.
Contact: Wing RA
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: TAATACGACTCCTATAGGG
High quality sequence stop: 838.
Location/Qualifiers
    1..880
    /organism="Gossypium arboreum"
    /mol_type="mRNA"
    /strain="AKA"
    /cultivar="8400"
    /db_xref="taxon:29729"
    /clone="GA_EB0002D19f"
    /tissue_type="Fibers isolated from bolls harvested 7-10
    dpa"
    /lab_host="E. coli"
    /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
    /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI"

BASE COUNT      249 a   176 c   176 g   278 t
ORIGIN

Query Match      25.9%; Score 494.6; DB 10; Length 880;
Best Local Similarity 74.6%; Pred. No. 2.9e-39;
Matches 634; Conservative 0; Mismatches 215; Indels 1; Gaps 1;

QY 778 AGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 837
Db 1 AGTGAATGGGACCAACTTGCAGTTGGATAGTGAACATGAATTCACAGTGGAAATGT 60

QY 838 GTCTGGCTGGTTTCAGATTCACAGCTTATATAATGTGTACAGAAATGGGTATCGTTACA 897
Db 61 GTTTGGTTAATCCAGTTACACAGATTGTATATGTGTACAGCAATGGGAATGTAAAA 120

QY 898 TCATTCCAAATCTCTTGACAACAATTTTCGTTCTCTCTTTTGAGGTACTATTGATCCA 957
Db 121 TCTTTTCAGATAATCTCGACAATGTTTTCATTCCTCACTATTTTGAAGTTACAGTTGATCCA 180

QY 958 GCTTCACACCCACAGCTCCATGCTCTCTGAGCAGGTGTGAGGTTGGACCTGGTTGAT 1017
Db 181 AATTCTCATCTCAACTACAGCTGTTCTTAAAAATGGTGGTTGGCTTTGATTTAGTTGAT 240

QY 1018 GATGAAGTAAACAGAAAGCGTCCAAACAAAGCACATGCCACACCTGAAACAGTGGAC 1077
Db 241 GATGAGATTAACAGAGAGCGCTCTTACAAAGCACATGCCAATCTCTGCTGAATGACT 300

QY 1078 AATGTTTCAACCTCGCATTTTCATATATATGCGTACTGCTATGCTATGCTTATTCACC 1137

```

```
Db 301 AATGAATTTAATCCTGCATATCTTCTTATTATGCGCTACTACTTTTATGCAAAACCTTTATACA 360
Qy 1138 CTAACAAGCTGGTGGTCAAGGGAATGACACATCATCAATTCGGTCCACATGCTGA 1197
Db 361 CTAACAAGCTGGTGGTCAAGGGAATGACACATCAATTCGGTCCACATGCTGA 420
Qy 1198 GAGGCTGGAGATGTTGATCAGTTCGAGGAGACATTTCTTCTCTGTGCACACATATCACAT 1257
Db 421 GAGGCTGGTGGATGATGACCATTTAGCTGCTGCTTCTGTGCAACAATATATCTCAC 480
Qy 1258 GGAATTAATCTAAGGAAGTCTCTGTGCTTCAAGTACTGTATCTTGGTGCAGATTGCT 1317
Db 481 GGAATTAATCTCGGAATCCCTGTTTGGCAGTACTGTATTAACCTCGCTCAGATCGGA 540
Qy 1318 CTGGGATCTCCCATTTGAGCAACACTCTTATTTCTTGTGACTATCATCGCAACCTTTT 1377
Db 541 TTGGCCATGCTCTCTGAGCAATAATTCCTTTCTTGAGTATCATCGCAACCAATTT 600
Qy 1378 CCAACGTTCTTCAACGAGTCTGAATGTCTCATTTATCTACGGATGACCCCTTTGCAAAAT 1437
Db 601 CTTTCAATTTTTCAGCTGGGTAATGTTCTTCTTCTATCTGATGACCCCTTTACAAAT 660
Qy 1438 CACCTGACAAAAGAACATTTGGTGAAGATACAGACATTTGCTGCTTGGTGGAGCTC 1497
Db 661 CATTTGACGAAGGAGCGCTTGTGGAAGATACAGTGTTCGACACAGGTTTGGAGCTC 720
Qy 1498 AGTTCTGTGATTTATGCGAATTCGAGGAACTCTGTTTACCAATCTGGTTTTCACAT 1557
Db 721 AGTCTGTGACTTATGTGAGATGCTAGAAATCTGTATACCAAGTCTGGGTTTTCAC 780
Qy 1558 GCTCTCAAGCGCACCT-GGATTTGTAAGAACTACTTTCAAAAGAGGACCTGCTGGAATGA 1616
Db 781 ATGTCAAGTTGATGCTGGCTTGGCAACAATATTTCTTAAGAGGCTCCAGAGGATGA 840
Qy 1617 TATTCACAGA 1626
Db 841 CANTCAAAAA 850

RESULT 14
BG365787
LOCUS
DEFINITION
913 bp mRNA linear EST 22-OCT-2001
HVSME10004D04f Hordeum vulgare 20 DAP spike EST library HVCNDA0010
(20 DAP) Hordeum vulgare subsp. vulgare cDNA clone HVSME10004D04f,
mRNA sequence.
BG365787
BG365787.2 GI:16323876
EST.
Hordeum vulgare subsp. vulgare
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Triticeae; Hordeum.
1 (bases 1 to 913)
Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Begum, D., Frisch, D., Yu
.Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi, D.W., Fenton
.R.D., Close, S.J., Oates, R. and Main, D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex 20 DAP spike cDNA library
Unpublished
On Mar 8, 2001 this sequence version replaced gi:13254886.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 406
Seq primer: AATAACCTCTACTAAAGG
High quality sequence start: 2
High quality sequence stop: 622.
Location/Qualifiers
```

## source

```
1. 913
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/db_xref="taxon:112509"
/clone="HVSME10004D04f"
/tissue_type="20 DAP spike"
/lab_host="SOLR"
/clone_lib="Hordeum vulgare 20 DAP spike EST library
HVCNDA0010 (20 DAP)"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spikes with awns trimmed were collected at 20 DAP (Fenton
). Total RNA was prepared, poly(A) RNA was purified, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagemids in the TJ Close lab at the University of
California, Riverside (Choi). Phagemids were plated and
picked at the Clemson University Genomics Institute (CUGI)
(Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA
preparations, DNA sequencing and sequence analysis were
performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates
, Rambo, Main). The sequence has been trimmed to remove
vector sequence and contains a minimum of 100 bases of
phred value 20 or above. For more details on library
preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders/Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
```

BASE COUNT 274 a 189 c 187 g 262 t  
ORIGIN

Query Match 25.7%; Score 490.6; DB 10; Length 913;  
Best Local Similarity 84.3%; Pred. No. 6.9e-39;  
Matches 575; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

Qy 12 GGTCTCTCATGCGGAGAGAGGTGATTAATGACCCCTGTACTCCAAAACCTAACCCAA 71  
Db 232 GGTCTCTCATGCGGAGAGAGGTGATTAATGACCCCTGTACTCCAAAACCTAACCTAA 291  
Qy 72 CCGCTTCACTTATGTCCTGAACAAAGTCAGAGCATGTTTCCAAACTGTTGATGGGT 131  
Db 292 CCCATTCATTTCTGTCCTGAACAAATCGAGCATGTTTCCAAATGTTGATGGGT 351  
Qy 132 TATCATGTTTATGCGGATAAAGATTGTACGGAGAGCATTTATCTGTGGCTGATGCTAC 191  
Db 352 TGTCCAGGTCTATGCGGATAAAGACTGTACGGAAAGGATTTATCTGTGCTGATGCTAC 411  
Qy 192 AACCTTCTCACTGACTTGCTGATTAATTTCTCCGAGTAACGGCTGAGGGAACACAGAAC 251  
Db 412 AACCTTTTTCACCGACTTGCTGATTAATTTCTCCGAGTAACGGCTGAGGGAACACAGAAC 471  
Qy 252 TGCTGCGCATATCCGTTAAATCTTCTTGAGCATAGTTAAATTCATCTCATGTTTAA 311  
Db 472 TGCTGTCATTAACCCGGTTGAATCTTCTAGAGCATAGTTTCAATTTTCATGATGTTTAA 531  
Qy 312 TCGCGATAGGGAATTTCTTTGCCAGAGAAGCTGCCCCACATCGTGATTTTTTACAATGTCAG 371  
Db 532 CGCGGACAGGAGTTTCTTGGCAGAGACTGCACCACATCGTGATTTTACAATGTTAG 591  
Qy 372 GAAGTTTGACACTCATGTTTCATCTATCAGCATGATGATCAAAACATCTCTGTGAGTT 431  
Db 592 GAAAGTCGACACTCATGTTTCACCTCAGCATGATGAATCAGAAACATTTTCTGAGATT 651  
Qy 432 CATAAATCCAAACTTAAGAAAGAACCTTGATGAGTGGTTCATTTTCAGAGATGGTACTTA 491  
Db 652 CATCAAGTCCAAACTTTGAGAAAGAACCTTGATGAGTGGTTCATTTTCAGAGATGGTACTTA 711  
Qy 492 TATG-ACITTTAAAGAGGTTTTTTGAGAGCTTGAAGCTTAACCTGGGTATGATCTCTGAATGTTG 550

```

Db      712 TATGACTTTAAAGGAGGTTTGGAGCTTGGACTTAACCTGGTATGACTGGAATGTTG 771
Qy      551 ATTGCTAGATGTCATGACGACAAAGACATTTTC-ATCGTTTGGACAAATTCATCTTA 609
Db      772 ATTGCTAAATGTCACGCTGACAAAAGTACTTTTCAATCGTTTGAACAAATCAAACCTT 831
Qy      610 AATACAAATCCATGTCGCAAAAGTAGGCTCAGAGAAATTTTCTCAACAAAGATATCTT 669
Db      832 AATACCAATCCATGTCGCGCAAGTTAGTTACGGGAAATTTTCTCTAAACGGGACATCT 891
Qy      670 ATTCAAGGCGTTTCTTCTGCTG 691
Db      892 TATTAAGGCAATTTTGGCTG 913

```

```

RESULT 15
LOCUS   BG442858
DEFINITION GA_Ea001803f Gossypium arboreum 7-10 dpa fiber library Gossypium
ARBOREUM cDNA clone GA_Ea001803f, mRNA sequence.
ACCESSION BG442858
VERSION   BG442858.1 GI:13352510
KEYWORDS EST.
SOURCE   Gossypium arboreum
ORGANISM Gossypium arboreum

```

```

REFERENCE 1 (bases 1 to 909)
AUTHORS  Wing,R.A., Frisch,D., Yu,Y., Main,D., Rambo,T., Simmons,J., Henry
          D., Wood,F.C., Leslie,A. and Wilkins,T.A.
TITLE     An integrated analysis of the genetics, development, and evolution
          of the cotton fiber
JOURNAL   Unpublished
COMMENT   Contact: Wing RA
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 7288
          Fax: 864 656 4293
          Email: rwing@clemson.edu
          Seq primer: TAATACGACTCACTATAGG
          High quality sequence stop: 714.
          Location/Qualifiers
            1..909
              /organism="Gossypium arboreum"
              /mol_type="mRNA"
              /strain="AKA"
              /cultivar="8400"
              /db_xref="taxon:29729"
              /clone="GA_Ea001803f"
              /tissue_type="fibers isolated from bolls harvested 7-10
              dpa"
              /lab_host="E. coli"
              /clone_lib="Gossypium arboreum 7-10 dpa fiber library"
              /note="Vector: pBK-CMV; Site 1: EcoRI; Site 2: XhoI"

```

## FEATURES

```

source
1..909
Query Match      25.6%; Score 488.2; DB 10; Length 909;
Best Local Similarity 75.4%; Pred. No. 1.2e-38;
Matches 673; Conservative 0; Mismatches 213; Indels 7; Gaps 5;

```

```

Qy      574 AAAGCAGATTTTCATCGTTTGGACAAATTCATCTAAATACATCCATGGCCAAAGT 633
Db      2  AAGAGCAGCTTTCATCGATTTGACAAATTCATCTAAATATAATCTTGTGGGCAAGC 61
Qy      634 AGGCTCAGAGAAATTTTCTCAACAAAGATATCTTATTCAAGCGCGTTTCTTCTGCTGAG 693
Db      62  AGACTTAGAGATCTTTTAAAGCAGGACATCTTATCCAGGACGTTTCTGCGAGAA 121
Qy      694 TTGACAAAGCAAGTTTCTCTGACCTTTCTGCTAGCAAAATATCAGATGGCAGAAATATAGG 753

```

```

Db      122 GTTCAAAGCAGTGTCTATCAGATCTTGAACAAGCAAAATATCAGATGGCAGAGTACAGG 181
Qy      754 ATTTCAATCTACGGAAGGAAAACAGAGTGAATGGGACCAAACTTGGAAAGTTGGATAGTGAAC 813
Db      182 GTGTCATATATGGAAGGAAAACAAAGTGAATGGGACCAAGTTGGCAGTTGGTTATTATTAAC 241
Qy      814 AATGAATGGCAGTGGGAAATGTTGTCTGGCTGGTTTCAGATTCACCGCTTATATAATGTG 873
Db      242 AATGAAATTTTATAGTGAGACTACTGTTTGGTTAATCCAGTTTACCACGATTTGATAATGTG 301
Qy      874 TACAAGGAAATGGGTATCGTTTACATCATTTCCAAATCTTCTTGACAAATTTTCGTTCTT 933
Db      302 TACAAGCAAAATGGGAATTTGTAATAATCTTTTCAGAAATATTTCTGGCAATATGTTTCAATCCA 361
Qy      934 CTTTTTGAAGTTTACTATTGATCCAGCTTTCACACCCACAGCTCCCATGTCTTCTCTGAAGCAG 993
Db      362 CTATTTGAAGTTACAGTTTGATCCAAATTTCTCATCTCAACTACACGTGTTCTCTAAAGATG 421
Qy      994 GTTGTAGGTTGGACCTGGTTGATGATGAAGTAAACCCAGAAAGCGTCCCAACAAAGCAC 1053
Db      422 GTGGTTGGCTTTGATTTAGTTGATGATGAGAGTAAACCCAGAGAGCGCTCTACAAAGCAC 481
Qy      1054 ATGCCCAACCTGACAGTGGACCAATGTGTTCAACCTGCAATTTTCATATATATGCGTAC 1113
Db      482 ATGCCAACTCTGCTGAATGGACTAATGAATTTAATCTCGCATATTTCTTATATGCTCTAC 541
Qy      1114 TACTGCTATGCTAACTTATTTCACCTTAAACAGCTGGTGAAGTCAAGGGGAATGACCACT 1173
Db      542 TACTTTTATGCAAACTTTTATACACTCAACAGCTTCTGAGTCAAAAGGAAATGCAGACA 601
Qy      1174 ATCAAATTCCTCCACATGCTGGAGAGCTGGAGATGTTGATCACTTGGCAGCGACATTT 1233
Db      602 ATAAACTCCGACCTCACTGTGGGAGGCTGGTGATATTGACCATTTAGCTGCTGCTTTC 661
Qy      1234 CTTCTCTGTCAACATATACATGGAATTAATCTAAGGAAGTCTCTGCTGCTTCAGTAC 1293
Db      662 CTTCTGTG-CACAATATATCTCACGGGATTAATCTCCGGAATCCCTGTTTTCAGTAC 720
Qy      1294 TTGTACTATCTTGGTTCAGATTGGTCTGGCGATGTCCTCCCATTTGAGCAACAACCTCTTATTT 1353
Db      721 TTGTATTACCTCGCTCAGATCGGATTTGGCCATGTCCTCTGAGCAATAATTCCTTTTG 780
Qy      1354 CTTGACTATCATCGCAACCCCTTTTCCAAAGTCTTCCAAAGAGGTCTGAATGTCTCATTA 1413
Db      781 CTGGACTTTTATTGG--ACCCCTTTCTTCAATCTTT-AAACGTGGGCTAAATGTTGCTCTT 837
Qy      1414 TCTACGGATGACCTTTTGCATAATTCACCTGACAAAGAACCAATTTGGTGGGAAGA 1466
Db      838 TCT-CTGATGACCTTTTACAA--TTCTTTGACAAAGAGAACCTTTTGGGGAATA 887

```

Search completed: December 15, 2003, 21:05:43  
Job time : 3676 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:20:48 ; Search time 46 Seconds  
(without alignments)  
1994.434 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRVRVAPWEKEVINDPCTPK.....NEMKLIVSYDNEILIPDELIDL 578

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A: Geneseq 19Jun03.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3100	100.0	578	22	AA1948
2	3082	99.4	603	22	AA1954
3	2683.5	86.6	681	22	AA1955
4	2644.5	85.3	839	20	AA19456
5	2644.5	85.3	839	23	AA19456
6	2644.5	85.3	859	20	AA19457
7	2342	75.5	595	22	AA1956
8	2338.5	75.4	492	22	AA1949
9	1706.5	55.0	345	22	AA1957

10	1613	52.0	658	22	AB865426	Drosophila melanog
11	1502.5	48.5	747	23	ABG70650	Rabbit adenosine m
12	1497	48.3	1813	22	ABG02232	Novel human diagn
13	1357.5	43.8	292	23	AAO16944	A thaliana AMP dea
14	627	20.2	888	22	AAO16944	S cerevisiae apopt
15	404	13.0	82	22	AA1950	Amino acid sequenc
16	344.5	11.1	107	22	AA1951	Amino acid sequenc
17	293	9.5	100	23	ABP09592	Human ORFX protein
18	191.5	6.2	160	22	AA1951	Human colon cancer
19	184	5.9	65	22	ABG50907	Human liver peptid
20	184	5.9	65	22	ABG30880	Human liver peptid
21	184	5.9	65	22	AAW56948	Human brain expres
22	184	5.9	65	22	AAW69232	Human bone marrow
23	184	5.9	65	23	ABG38841	Human peptide enco
24	142	4.6	2138	23	AAE25295	Human nucleic acid
25	141	4.5	293	22	ABG12977	Novel human diagn
26	134	4.3	553	22	ABG61988	Drosophila melanog
27	126	4.1	467	22	AB8711743	Drosophila melanog
28	120.5	3.9	352	23	AB853593	Lactococcus lactis
29	113.5	3.7	370	22	AB864809	Drosophila melanog
30	111	3.6	1498	22	AB864857	Drosophila melanog
31	111	3.6	1498	23	AAE25146	Fruit fly ARMS pro
32	107	3.5	782	11	AAO69991	Polypeptide antige
33	106.5	3.4	999	18	AAW26682	Bovine lysosomal a
34	106.5	3.4	999	18	AAW26684	Bovine lysosomal a
35	106.5	3.4	1032	23	AAO21588	Murine Toll-like r
36	105	3.4	464	13	AAE29651	AmEPV Sporoidein a
37	105	3.4	464	19	AAW41297	AmEPV entomopoxv
38	105	3.4	464	20	AAW30165	Protein encoded by
39	105	3.4	464	23	ABJ10372	Amino acid encoded
40	105	3.4	464	23	AB877627	AmEPV core protein
41	105	3.4	627	23	AB855152	Lactococcus lactis
42	104.5	3.4	511	19	AAW70287	Dendritic cell-der
43	104.5	3.4	667	23	AB892524	Herbicidally activ
44	104.5	3.4	773	22	AB866567	Drosophila melanog
45	104.5	3.4	773	22	AB870352	Drosophila melanog

ALIGNMENTS

RESULT 1

AA1948

ID AA1948 standard; Protein; 578 AA.

XX

AC AA1948;

XX

DT 15-MAY-2001 (first entry)

XX

DE Amino acid sequence of a corn AMP deaminase enzyme.

XX

DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;

XX

KW inosine.

XX

OS Zea mays.

XX

PN WO200109305-A2.

XX

PD 08-FEB-2001.

XX

PF 28-JUL-2000; 2000MO-US21009.

XX

PR 30-JUL-1999; 99US-0146473.

XX

PA (DUPO) DU PONT DE NEMOURS & CO E I.

XX

PI (PION-) PIONEER HI-BRED INT INC.

XX

PI Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

XX

DR WPI; 2001-159866/16.

XX

DR N-PSDB; AAF25488.

XX

PT New polynucleotides encoding AMP deaminase or adenosine deaminase used

PT e.g. for designing or identifying herbicides that inhibit the enzyme  
XX activities, and as probes for genetic or physical mapping -  
PS Claim 10; Page 43-44; 72pp; English.

XX The present sequence represents an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

XX Sequence 578 AA;  
Query Match 100.0%; Score 3100; DB 22; Length 578;  
Best Local Similarity 100.0%; Pred. No. 1.3e-299;  
Matches 578; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PRVRAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 60  
DB 1 PRVRAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 60  
QY 61 ADATFTDHLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHRDF 120  
DB 61 ADATFTDHLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHRDF 120  
QY 121 YNVRKVDTHVHSACMNQKHLRFIKSKURKEPDEVVIFRDGTMTLKEVFESLDLTGYD 180  
DB 121 YNVRKVDTHVHSACMNQKHLRFIKSKURKEPDEVVIFRDGTMTLKEVFESLDLTGYD 180  
QY 181 LVNDDLVDHADKSTFHRFDKFNLYKPCQSRLREIFLQDNLIQGRFLAELTKQVFSDL 240  
DB 181 LVNDDLVDHADKSTFHRFDKFNLYKPCQSRLREIFLQDNLIQGRFLAELTKQVFSDL 240  
QY 241 SASKYQMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLQIPLRYNLYKEMGIVTS 300  
DB 241 SASKYQMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLQIPLRYNLYKEMGIVTS 300  
QY 301 FQNLNLI FVPLFEVITDPAHPOLHVLKQVGLDLDVDESKPERRPKHMPTPEQWTN 360  
DB 301 FQNLNLI FVPLFEVITDPAHPOLHVLKQVGLDLDVDESKPERRPKHMPTPEQWTN 360  
QY 361 VFNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLCHNISHG 420  
DB 361 VFNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGVDVHLAATFLCHNISHG 420  
QY 421 INLRKSPVLYLYLQGIQGLAMSPLNNSLFLDYHRNPFTTFFQGLNVLSTDDPLQIH 480  
DB 421 INLRKSPVLYLYLQGIQGLAMSPLNNSLFLDYHRNPFTTFFQGLNVLSTDDPLQIH 480  
QY 481 LTKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWTGKNYKRGAGNDI 540  
DB 481 LTKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWTGKNYKRGAGNDI 540  
QY 541 HRTNVPHIRVQPREMIWRNEMKLVYSDNEILLIPDELDL 578  
DB 541 HRTNVPHIRVQPREMIWRNEMKLVYSDNEILLIPDELDL 578

RESULT 2  
AAB31954

ID AAB31954 standard; Protein; 603 AA.  
XX AAB31954;  
XX 15-MAY-2001 (first entry)  
DT Amino acid sequence of a soybean AMP deaminase enzyme.  
DE AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
KW inosine.  
XX Glycine max.  
XX Key Location/Qualifiers  
FT Misc-difference 15 /note= "unknown residue encoded by CAG"  
FT  
XX WO200109305-A2.  
XX 08-FEB-2001.  
XX 28-JUL-2000; 2000WO-US21009.  
XX 30-JUL-1999; 99US-0146473.  
XX (DUPO ) DU PONT DE NEMOURS & CO E I.  
XX (PION-) PIONEER HI-BRED INT INC.  
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;  
XX WPI; 2001-159866/16.  
XX N-PSDB; AAF25494.  
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used  
PT e.g. for designing or identifying herbicides that inhibit the enzyme  
PT activities, and as probes for genetic or physical mapping -  
XX Claim 10; Page 57-59; 72pp; English.  
XX The present sequence represents an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of  
CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

QY 2 RVRVAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 61  
DB 27 REEVAPWEKEVINDCTPKPNPFTYVPEPKSEHVFTQDGVHIVYADKCTESIYPV 86  
QY 62 DATTEFTDHLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHRDF 121  
DB 87 DATTEFTDHLHYLTRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQTAHRDF 146  
QY 122 NVVRKVDTHVHSACMNQKHLRFIKSKURKEPDEVVIFRDGTMTLKEVFESLDLTGYD 181

Query Match 99.4%; Score 3082; DB 22; Length 603;  
Best Local Similarity 99.7%; Pred. No. 8.8e-298;  
Matches 575; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



Db 147 NVRKVDTHVHHSACNMQRHLLRFKSKLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 206  
QY 182 NVDLLDVHADKSTFHRDFKFNLYKYNPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 241  
Db 207 NVDLLDVHADKSTFHRDFKFNLYKYNPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 266  
QY 242 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNNVYKEMGIVTSF 301  
Db 267 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNNVYKEMGIVTSF 326  
QY 302 QNLLDNIFVLPFEVITDPASHPOLHVLKQVGLDLVDDSEKPERPRTKHMPTPQWNTV 361  
Db 327 QNLLDNIFVLPFEVITDPASHPOLHVLKQVGLDLVDDSEKPERPRTKHMPTPQWNTV 386  
QY 362 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGDVVDHLAATFLLCHNISHGI 421  
Db 387 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGDVVDHLAATFLLCHNISHGI 446  
QY 422 NLRKSPVLQYLYLQIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSTDDPLQIHL 481  
Db 447 NLRKSPVLQYLYLQIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSTDDPLQIHL 506  
QY 482 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKYNFKRGPAGNDIH 541  
Db 507 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKYNFKRGPAGNDIH 566  
QY 542 RTNVPHIRVQFREMWRNEMKLVSDNEILIPDEL 578  
Db 567 RTNVPHIRVQFREMWRNEMKLVSDNEILIPDEL 603

RESULT 3

AAB31955  
ID AAB31955 standard; Protein; 681 AA.

XX  
AC AAB31955;

XX 15-MAY-2001 (first entry)

XX Amino acid sequence of a wheat AMP deaminase enzyme.

XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
KW inosine.

XX Triticum aestivum.

XX WO200109305-A2.

XX 08-FEB-2001.

XX 28-JUL-2000; 2000WO-US21009.

XX 30-JUL-1999; 99US-0146473.

XX (DUPO ) DU PONT DE NEMOURS & CO E I.

XX (PION-) PIONEER HI-BRED INT INC.

XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;

XX WPI; 2001-159866/16.

XX N-PSDB; AAF25495.

XX New polynucleotides encoding AMP deaminase or adenosine deaminase used  
PT e.g. for designing or identifying herbicides that inhibit the enzyme  
PT activities, and as probes for genetic or physical mapping

XX Claim 10; Page 60-62; 72pp; English.

XX The present sequence represents an AMP deaminase. The specification also  
CC describes adenosine deaminase. These enzymes convert adenosine to  
CC inosine. Mutations in these genes cause disruptions in then salvage and  
CC catabolism of adenosine and AMP. In humans, this may lead to death of

CC white blood cells, which causes severe immunodeficiencies. The AMP  
CC deaminase and adenosine deaminase may be used to prepare antibodies  
CC to these proteins, and to design or identify herbicides that inhibit  
CC their enzyme activities. The polynucleotides are used as probes for  
CC genetically and physically mapping genes that they compose, and as  
CC markers for traits linked to those genes, where such information may be  
CC used in plant breeding to develop lines with desired phenotypes. The  
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding  
CC homologous proteins from the same or other plant species, and in  
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic  
CC acid fragments encoding homologous genes from DNA or RNA. These may  
CC also be used to create transgenic plants in which the polypeptides are  
CC overexpressed or suppressed, and as probes in direct fluorescent in  
CC situ hybridisation (FISH).

XX  
SQ Sequence 681 AA;

Query Match 86.6%; Score 2683.5; DB 22; Length 681;

Best Local Similarity 85.1%; Pred. No. 5.6e-250;

Matches 490; Conservative 42; Mismatches 43; Indels 1; Gaps 1;

QY 2 RVRVAPWEKEVINDPCTPKPNPFTYVPEKSEHVFOVDGVHIVYADKDCETSIYPA 61  
Db 106 REEVAPWEKEIITDSTPKPNPFTYEQQTKTEHFMVDGVHIVYFNKAKERIYPA 165

QY 62 DATTFPTDLHYLRVTAAGNTRTVCHNRNLNLEHKKFHLMLNADREFLAQKTAPHRDFY 121  
Db 166 DATTFPTDMVILRVLAAGDIRTVCYKRLNLEQKFNHLMLNADRELLAQAAPHRDFY 225

QY 122 NVRKVDTHVHHSACNMQRHLLRFKSKLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 181  
Db 226 NVRKVDTHVHHSACNMQRHLLRFKSKLRKEPDEVIFRDGYMTLKEVFSLDLTGYDL 285

QY 182 NVDLLDVHADKSTFHRDFKFNLYKYNPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 241  
Db 286 NVDLLDVHADKSTFHRDFKFNLYKYNPCGQSRLEIFLKQDNLIQGRFLAELTKQVFSDL 345

QY 242 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNNVYKEMGIVTSF 301  
Db 346 ASKYQMAEYRISYGRKQSEWDLASWVNNELHSGNVVWLQVLPRLYNNVYKEMGIVTSF 405

QY 302 QNLLDNIFVLPFEVITDPASHPOLHVLKQVGLDLVDDSEKPERPRTKHMPTPQWNTV 361  
Db 406 QNLLDNIFVLPFEVITDPASHPOLHVLKQVGLDLVDDSEKPERPRTKHMPTPQWNTV 465

QY 362 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGDVVDHLAATFLLCHNISHGI 421  
Db 466 FNPAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGDVVDHLAATFLLCHNISHGI 525

QY 422 NLRKSPVLQYLYLQIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSTDDPLQIHL 481  
Db 526 NLRKSPVLQYLYLQIGLAMSPLSNNSLFLDYHRNPFTTQGRGLNVSLSTDDPLQIHL 585

QY 482 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKYNFKRGPAGNDIH 541  
Db 586 TKEPLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKYNFKRGPAGNDIH 645

QY 542 RTNVPHIRVQFREMWRNEMKLVSDNEILIPDEL 577  
Db 646 QTNVPHIRIEPRHTIWEEMELIHRN-VDIPEID 680

RESULT 4

AAY33456

ID AAY33456 standard; Protein; 839 AA.

XX  
AC AAY33456;

XX 13-DEC-1999 (first entry)

XX A. thaliana AMP-deaminase protein #1.

XX AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;

inhibitor; resistance.  
Arabidopsis thaliana.  
W09950400-Al.  
07-OCT-1999.  
25-MAR-1999; 99WO-EP02016.  
01-APR-1998; 98DE-1014512.  
(BADI ) BASF AG.  
Lerchl J, Reindl A;  
WPI; 1999-580759/49.  
DR N-PSDB; AA223395.  
DNA encoding AMP deaminase, assay systems for identifying inhibitors and transgenic plants -  
Claim 4; Page 34-37; 45pp; German.  
This invention describes a novel Arabidopsis thaliana AMP-deaminase (adenosine triphosphate aminohydrolase, EC 3.5.4.6) which has herbicidal activity. The AMP deaminase DNA sequence, can be introduced into pro- or eukaryotic cells, with the relevant control elements to control transcription and translation in the cell. An expression cassette derived from the products of the invention is used to transform plants. It is useful for producing a test system to identify inhibitors of AMP deaminase. Plants containing the expression cassette are useful for the production of the AMP deaminase. The plants have an increased resistance against inhibitors of the AMP-deaminase through strengthened expression of the AMP deaminase DNA sequence. The expression cassette is useful for producing plants with an increased content of AMP. This sequence represents the A. thaliana AMP-deaminase described in the invention.  
Sequence 839 AA;  
Query Match 85.3%; Score 2644.5; DB 20; Length 839;  
Best Local Similarity 83.9%; Pred. No. 6e-254;  
Matches 480; Conservative 46; Mismatches 45; Indels 1; Gaps 1;  
QY 5 VAPWEKEVINDPCTPKPNPFTYVPEPKSEHVQTVGVHVVYADKCTESIYPVADAT 64  
DB 267 VAPWEKEVISDPSTPKNTEPFAHYPOGKSDHCFEMQDGVVHVFAKDAKEDLFPVADAT 326  
QY 65 TFFTDLHYLRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQKTAPHRDPYNNR 124  
DB 327 AFTDLHLVVKVTAAGNIRTLCHRLVLLLEQENLHMLNADKEFLAQSAPHRDPYNNR 386  
QY 125 KVDTHVHHSACMNQKHLRFIKSLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD 184  
DB 387 KVDTHVHHSACMNQKHLRFIKSLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD 446  
QY 185 LLDVHADKSTFRHFDKFNLYKPCGSRRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244  
DB 447 LLDVHADKSTFRHFDKFNLYKPCGSRRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 506  
QY 245 YQMAEYRISYGRKQSEWQLASWIVNNLHSGNVVWLQVLPRLVNVYKEMGIVTSFQNL 304  
DB 507 YQMAEYRISYGRKSEWQLASWIVNNLHSGNVVWLQVLPRLVNVYKEMGIVTSFQNI 566  
QY 305 LQNFVPLFEVITDAPSHQVFLKQVGLDLVDDSEKPERRPRTKHMPTPQWTVNFP 364  
DB 567 LQNFIFLFEATVDPDSHPQVFLKQVGLDLVDDSEKPERRPRTKHMPTPQWTVNFP 626  
QY 365 AFSYVYVCYANLFTLNKLRSEKGMTTIKFRPHAGEAGVDVHLAATFLCHNISHGINLR 424  
DB 627 AFSYVYVCYANLVNKLKRESKGMTTITLRPHSGEAGDIDHLAAFTFLCHSHGINLR 686  
QY 425 KSPVQLYYLQIGLAWSPLSNNSFLDYHRNPPFPVFFLRGLNVSLSLTDPLQIHLTKE 484  
DB 687 KSPVQLYYLQIGLAWSPLSNNSFLDYHRNPPFPVFFLRGLNVSLSLTDPLQIHLTKE 746  
QY 485 PLVEYSSTAASLWKLSSCDLCEIARNVSQSGFHALKAHWIGKNYFKRGGAGNDIHRTN 544  
DB 747 PLVEYSSTAASVWKLKACDLCEIARNVSQSGFHALKSHWIGKDYKRGDPGNDIHKTN 806  
QY 545 VPHIRVQFREMIRWNRNEMKLVYSDNEILIPDEL 576  
DB 807 VPHIRVEFRDTIWKEEMQOQVYL-GKAVISDEV 837  
RESULT 5  
AAO16943  
ID AAO16943 standard; Protein; 839 AA.  
XX  
AC AAO16943;  
XX  
DT 16-MAY-2002 (first entry)  
XX  
DE A thaliana AMP deaminase.  
XX  
KW AMP deaminase; adenosine monophosphate; transgenic plant;  
KW herbicide resistance; herbicide; inhibitor.  
XX  
OS Arabidopsis thaliana.  
XX  
FN WO200206319-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 06-JUL-2001; 2001WO-EP07767.  
XX  
PR 17-JUL-2000; 2000DE-1035084.  
XX  
PA (AVET ) AVENTIS CROPS SCIENCE GMBH.  
XX  
PI Schulz A, Streiber W, Hanke C, Schmidt F, Schubel A;  
XX  
DR WPI; 2002-195802/25.  
XX  
DR N-PSDB; AAL45016.  
XX  
PT New nucleic acid for plant adenosine monophosphate deaminase, useful in screening for herbicides and preparing herbicide-resistant plants -  
XX  
PS Claim 1; Page 44-47; 51pp; German.  
XX  
CC The present invention provides the protein and coding sequences of the Arabidopsis thaliana adenosine monophosphate (AMP) deaminase. The coding sequence can be used to transform prokaryotic or eukaryotic cells, especially to produce transgenic plants (e.g. barley, rice, soya etc.) resistant to herbicidal inhibitors of AMP-deaminase, for recombinant production of proteins with AMP-deaminase activity, and to identify related genes in other organisms. The protein can be used for identification and biochemical/structural characterisation of new AMP-deaminase inhibitors and potential herbicides. The present sequence is the protein of the invention.  
XX  
SQ Sequence 839 AA;  
Query Match 85.3%; Score 2644.5; DB 23; Length 839;  
Best Local Similarity 83.9%; Pred. No. 6e-254;  
Matches 480; Conservative 46; Mismatches 45; Indels 1; Gaps 1;  
QY 5 VAPWEKEVINPCTPKPNPFTYVPEPKSEHVQTVGVHVVYADKCTESIYPVADAT 64  
DB 267 VAPWEKEVISDPSTPKNTEPFAHYPOGKSDHCFEMQDGVVHVFAKDAKEDLFPVADAT 326  
QY 65 TFFTDLHYLRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQKTAPHRDPYNNR 124  
DB 327 AFTDLHLVVKVTAAGNIRTLCHRLVLLLEQENLHMLNADKEFLAQSAPHRDPYNNR 386  
QY 125 KVDTHVHHSACMNQKHLRFIKSLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD 184  
DB 387 KVDTHVHHSACMNQKHLRFIKSLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVD 446  
QY 185 LLDVHADKSTFRHFDKFNLYKPCGSRRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244  
DB 447 LLDVHADKSTFRHFDKFNLYKPCGSRRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 506  
QY 245 YQMAEYRISYGRKQSEWQLASWIVNNLHSGNVVWLQVLPRLVNVYKEMGIVTSFQNL 304  
DB 507 YQMAEYRISYGRKSEWQLASWIVNNLHSGNVVWLQVLPRLVNVYKEMGIVTSFQNI 566  
QY 305 LQNFVPLFEVITDAPSHQVFLKQVGLDLVDDSEKPERRPRTKHMPTPQWTVNFP 364  
DB 567 LQNFIFLFEATVDPDSHPQVFLKQVGLDLVDDSEKPERRPRTKHMPTPQWTVNFP 626  
QY 365 AFSYVYVCYANLFTLNKLRSEKGMTTIKFRPHAGEAGVDVHLAATFLCHNISHGINLR 424  
DB 627 AFSYVYVCYANLVNKLKRESKGMTTITLRPHSGEAGDIDHLAAFTFLCHSHGINLR 686  
QY 425 KSPVQLYYLQIGLAWSPLSNNSFLDYHRNPPFPVFFLRGLNVSLSLTDPLQIHLTKE 484

Db 387 KVDTHVHSCACMKELLPKSKLRKEPDEVIFRDGTYLTREVFESDLTGVDLNDV 446  
 QY 185 LLDVHADKSTFHRDFKFNLYKYNPCGOSRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244  
 Db 447 LLDVHADKSTFHRDFKFNLYKYNPCGOSRLREIFLKQDNLIQGRFLGEITKQVFSLEASK 506  
 QY 245 YQMAEYRISYGRKQSEWDQLASWVNNELHSGNVVLMVQIPLRYNYKEMGIVTSFQNL 304  
 Db 507 YQMAEYRISYGRKQSEWDQLASWVNNELHSGNVVLMVQIPLRYNYKEMGIVTSFQNI 566  
 QY 305 LQNIFFVPLEVTIDPASHPOLHVFELKQVVGDLVDDESKPERPPTKHMPTPEQWTVNFP 364  
 Db 567 LQNIFFVPLEVTIDPASHPOLHVFELKQVVGDLVDDESKPERPPTKHMPTPEQWTVNFP 626  
 QY 365 AFSYVYVYCYANLFTLNKLRKSGMTTIFRPHAGEAGDVHDHLAATFLCHNISHGINLR 424  
 Db 627 AFSYVYVYCYANLYVNLKLRKSGMTTITLRPHSGEAGDIDLHAATFLCHNISHGINLR 686  
 QY 425 KSPVLOYLYLQIGLAMSPLNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTK 484  
 Db 687 KSPVLOYLYLQIGLAMSPLNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTK 746  
 QY 485 PLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPGAGNDIHRTN 544  
 Db 747 PLVEEYSIAASVWKLACDLCEIARNVYQSGFSHALKSHWIGKDYKRGPGDNDIHKTN 806  
 QY 545 VPHIRVOFREMIRNEMKLVYSDNEILIPDEL 576  
 Db 807 VPHIRVEFRDTIWKEEMQOQVYL-GRAVISDEV 837

## RESULT 6

AAY33457  
 ID AAY33457 standard; Protein; 859 AA.  
 AC AAY33457;

DT 13-DEC-1999 (first entry)  
 DE A. thaliana AMP-deaminase protein #2.

KW AMP-deaminase; adenosine triphosphate aminohydrolase; plant; herbicide;  
 KW inhibitor; resistance.

OS Arabidopsis thaliana.

XX WO9950400-A1.

XX 07-OCT-1999.

XX 25-MAR-1999; 99WO-BP02016.

XX 01-APR-1998; 98DE-1014512.

XX (BADI ) BASF AG.

XX Lerchl J, Reindl A;

XX WPI: 1999-580759/49.

XX N-PSDB; AA223395.

PT DNA encoding AMP deaminase, assay systems for identifying inhibitors

PT and transgenic plants -

XX Disclosure; Fig 2; 45pp; German.

CC This invention describes a novel Arabidopsis thaliana AMP-deaminase  
 CC (adenosine triphosphate aminohydrolase, EC 3.5.4.6) which has herbicidal  
 CC activity. The AMP deaminase DNA sequence, can be introduced into pro- or  
 CC eukaryotic cells, with the relevant control elements to control  
 CC transcription and translation in the cell. An expression cassette  
 CC derived from the products of the invention is used to transform plants.

CC It is useful for producing a test system to identify inhibitors of AMP  
 CC deaminase. Plants containing the expression cassette are useful for the  
 CC production of the AMP deaminase. The plants have an increased resistance  
 CC against inhibitors of the AMP-deaminase through strengthened expression  
 CC of the AMP deaminase DNA sequence. The expression cassette is useful for  
 CC producing plants with an increased content of AMP. This sequence  
 CC represents the A. thaliana AMP-deaminase described in the invention.

XX Sequence 859 AA;

Query Match 85.3%; Score 2644.5; DB 20; Length 859;

Best Local Similarity 83.9%; Pred. No. 6.2e-254;

Matches 480; Conservative 46; Mismatches 45; Indels 1; Gaps 1;

QY 5 VAPKEKEVINDPCTPKPNPNPFTYVPEPKSEHVFTQVDGVIHVADKCTESIVPAADAT 64  
 Db 287 VAPKEKEVISDPSITPKNTEPFAHYPOCKSDHCEMOMQGVVHVFAKDAKEDLFFPAADAT 346  
 QY 65 TFFTDLHVLRVTAAGNTRTVCHNRNLNLEHKKFHLMLNADREFLAQKTAAPHRDFYVNR 124  
 Db 347 AFTDHLHVLKVIAGNIRTLCHRRLLVLEQKFNHLMLNADKGFLOKSAAPHRDFYVNR 406  
 QY 125 KVDTHVHSCACMKOKHLLRFIKSLRKEPDEVIFRDGTYLTREVFESDLTGVDLNDV 184  
 Db 407 KVDTHVHSCACMKOKHLLRFIKSLRKEPDEVIFRDGTYLTREVFESDLTGVDLNDV 466  
 QY 185 LLDVHADKSTFHRDFKFNLYKYNPCGOSRLREIFLKQDNLIQGRFLAELTKQVFSLSASK 244  
 Db 467 LLDVHADKSTFHRDFKFNLYKYNPCGOSRLREIFLKQDNLIQGRFLGEITKQVFSLEASK 526  
 QY 245 YQMAEYRISYGRKQSEWDQLASWVNNELHSGNVVLMVQIPLRYNYKEMGIVTSFQNL 304  
 Db 527 YQMAEYRISYGRKQSEWDQLASWVNNELHSGNVVLMVQIPLRYNYKEMGIVTSFQNI 586  
 QY 305 LQNIFFVPLEVTIDPASHPOLHVFELKQVVGDLVDDESKPERPPTKHMPTPEQWTVNFP 364  
 Db 587 LQNIFFVPLEVTIDPASHPOLHVFELKQVVGDLVDDESKPERPPTKHMPTPEQWTVNFP 646  
 QY 365 AFSYVYVYCYANLFTLNKLRKSGMTTIFRPHAGEAGDVHDHLAATFLCHNISHGINLR 424  
 Db 647 AFSYVYVYCYANLYVNLKLRKSGMTTITLRPHSGEAGDIDLHAATFLCHNISHGINLR 706  
 QY 425 KSPVLOYLYLQIGLAMSPLNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTK 484  
 Db 707 KSPVLOYLYLQIGLAMSPLNNSLFLDYHRNPPTFFQRLGNVSLSTDDPLQIHLTK 766  
 QY 485 PLVEEYSIAASLWKLSSCDLCEIARNVYQSGFSHALKAHWIGKNYFKRGPGAGNDIHRTN 544  
 Db 767 PLVEEYSIAASVWKLACDLCEIARNVYQSGFSHALKSHWIGKDYKRGPGDNDIHKTN 826  
 QY 545 VPHIRVOFREMIRNEMKLVYSDNEILIPDEL 576  
 Db 827 VPHIRVEFRDTIWKEEMQOQVYL-GRAVISDEV 857

## RESULT 7

AAB31956

ID AAB31956 standard; Protein; 595 AA.

XX AAB31956;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a corn adenosine deaminase enzyme.

KW AMP deaminase; adenosine deaminase; adenosine; transgenic plant;  
 KW inosine.

XX Zea mays.

XX WO200109305-A2.

XX 08-FEB-2001.

```
XX 28-JUL-2000; 2000WO-US21009.
PF 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
PA Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
XX N-PSDB; AAF25496.
PT New polynucleotides encoding AMP deaminase or adenosine deaminase used
PT e.g. for designing or identifying herbicides that inhibit the enzyme
PT activities, and as probes for genetic or physical mapping -
XX Claim 10; Page 63-65; 72pp; English.
XX The present sequence represents adenosine deaminase. The specification
CC also describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX Sequence 595 AA;
SQ
Query Match 75.5%; Score 2342; DB 22; Length 595;
Best Local Similarity 87.7%; Pred. No. 5e-224;
Matches 429; Conservative 32; Mismatches 28; Indels 0; Gaps 0;
QY 2 RRVAVPEKEVINDPCTKPNPNFTYVPEPKSEHVQTVGVIHYVADKDCETSIYPA 61
DB 107 REAVAPWKEVISEDPSPTKPNPDFFLYIEGNSDHYFEMQDQVIRYVPRDAKEELFPVA 166
QY 62 DATFTFDLHYILRVTAAGNTRTVCHNRLNLLHKKFKFLMLNADREFLAQKTAHRDFY 121
DB 167 DATFTFDLHLLRVIAAGNIRYLCHRLNLLHKKFKFLMLNADREFLAQKSAHRDFY 226
QY 122 NVRKVDTHVHHSACMNQKHLRFIKSKRKEPDEVVIFRDGTGYMTLKEVPESLDLTGYDL 181
DB 227 NVRKVDTHVHHSACMNQKHLRFIKSKRKEPDEVVIFRDGTGYTLTLEEVFKSLDLSGYDL 286
QY 182 NVDLLDVHADKSTFHRDKFNLYNCPGQSRRLREIFLKQDNLQGRFLAELTKQVSDLS 241
DB 287 NVDLLDVHADKSTFHRDKFNLYNCPGQSRRLREIFLKQDNLQGRFLAELTKQVSDLS 346
QY 242 ASKYQMAEYRISYIGRKQSEWDLASWVNNELHSGNVWLVQIPRLXNYKEMGIVTSF 301
DB 347 ASKYQMAEYRISYIGRKQSEWDLASWVNNELHSGNVWLVQIPRLXNYKEMGIVTSF 406
QY 302 QNLLDNIFVPEFTVIDPASHPOLHVFLKQVGLDLVDDESKPERPRTKMTPTPEQWTVN 361
DB 407 QNMLDNIFELFEVTYVNPDSHPOLHVFLKQVGLDLVDDESKPERPRTKMTPTPEQWTVN 466
QY 362 FNPAFSYAYCYANLFLINKLRSGKMTIKRPHAGEAGDVNDHLAATFLCHNLSHGI 421
DB 467 FNPAFSYAYCYANLFLINKLRSGKMTIKRPHAGEAGDVNDHLAATFLTAHNTAHI 526
QY 422 NLRKSPVLQYLYLQIGLAMSPLSNNSFLDYHRNPFPPFELRGLNLSLTDPLQIHL 481
DB 527 NLRKSPVLQYLYLQIGLAMSPLSNNSFLDYHRNPFPPFELRGLNLSLTDPLQIHL 586
QY 482 TKEPLVEEY 490
DB 587 TKEPLVEEY 595
RESULT 8
AAB31949
ID AAB31949 standard; Protein; 492 AA.
XX AAB31949;
XX 15-MAY-2001 (first entry)
XX Amino acid sequence of a rice AMP deaminase enzyme.
XX AMP deaminase; adenosine deaminase; adenosine; transgenic plant;
XX inosine.
XX Oryza sativa.
XX WO200109305-A2.
XX 08-FEB-2001.
XX 28-JUL-2000; 2000WO-US21009.
XX 30-JUL-1999; 99US-0146473.
XX (DUPO ) DU PONT DE NEMOURS & CO E I.
XX (PION-) PIONEER HI-BRED INT INC.
XX Caspar T, Falco SC, Sakai H, Weng Z, Hu X;
XX WPI; 2001-159866/16.
XX N-PSDB; AAF25489.
XX New polynucleotides encoding AMP deaminase or adenosine deaminase used
XX e.g. for designing or identifying herbicides that inhibit the enzyme
XX activities, and as probes for genetic or physical mapping -
XX Claim 10; Page 45-47; 72pp; English.
XX The present sequence represents an AMP deaminase. The specification also
CC describes adenosine deaminase. These enzymes convert adenosine to
CC inosine. Mutations in these genes cause disruptions in then salvage and
CC catabolism of adenosine and AMP. In humans, this may lead to death of
CC white blood cells, which causes severe immunodeficiencies. The AMP
CC deaminase and adenosine deaminase may be used to prepare antibodies
CC to these proteins, and to design or identify herbicides that inhibit
CC their enzyme activities. The polynucleotides are used as probes for
CC genetically and physically mapping genes that they compose, and as
CC markers for traits linked to those genes, where such information may be
CC used in plant breeding to develop lines with desired phenotypes. The
CC nucleic acid fragments may be used to isolate cDNAs and genes encoding
CC homologous proteins from the same or other plant species, and in
CC polymerase chain reaction (PCR) protocols to amplify longer nucleic
CC acid fragments encoding homologous genes from DNA or RNA. These may
CC also be used to create transgenic plants in which the polypeptides are
CC overexpressed or suppressed, and as probes in direct fluorescent in
CC situ hybridisation (FISH).
XX Query 89 RNLLEHKKFKFLMLNADREFLAQKTAHRDFYVNRKVDTHVHHSACMNQKHLRFIKSK 148
DB 4 RNLLEHKKFKFLMLNADREFLAQKTAHRDFYVNRKVDTHVHHSACMNQKHLRFIKSK 63
```







```

XX 16-MAY-2002 (first entry)
XX A thaliana AMP deaminase EST encoded fragment.
XX AMP deaminase; adenosine monophosphate; transgenic plant; EST;
KW herbicide resistance; herbicide; inhibitor; expressed sequence tag.
XX Arabidopsis thaliana.
XX WO200206319-A2.
XX 24-JAN-2002.
XX 06-JUL-2001; 2001WO-EP07767.
XX 17-JUL-2000; 2000DE-1035084.
XX (AVET ) AVENTIS CROPS SCIENCE GMBH.
XX Schulz A, Streiber W, Hanke C, Schmidt F, Schubel A;
XX WPI; 2002-195802/25.
XX N-PSDB; AAL45031.
XX New nucleic acid for plant adenosine monophosphate deaminase, useful in
PT screening for herbicides and preparing herbicide-resistant plants -
XX Example 2; Page 23; 51pp; German.
XX The present invention provides the protein and coding sequences of the
CC Arabidopsis thaliana adenosine monophosphate (AMP) deaminase. The coding
CC sequence can be used to transform prokaryotic or eukaryotic cells,
CC especially to produce transgenic plants (e.g. barley, rice, soya etc.)
CC resistant to herbicidal inhibitors of AMP-deaminase, for recombinant
CC production of proteins with AMP-deaminase activity, and to identify
CC related genes in other organisms. The protein can be used for
CC identification and biochemical/structural characterisation of new
CC AMP-deaminase inhibitors and potential herbicides. The present sequence
CC is an EST encoded fragment of the protein of the invention.
XX SQ Sequence 292 AA;
Query Match 43.8%; Score 1357.5; DB 23; Length 292;
Best Local Similarity 84.2%; Pred. No. 2.2e-126;
Matches 245; Conservative 24; Mismatches 21; Indels 1; Gaps 1;
QY 286 PRLNYVYKEMGIVTSFQNLNDNIFVLPFEVITDPASHPQLHVLKQVGVGLDLDVDESKPE 345
DB 1 PRLNYVYKEMGIVTSFQNLNDNIFVLPFEVITDPASHPQLHVLKQVGVGLDLDVDESKPE 60
QY 346 RRPTKHMPTEQWNTVNFNPAFSYVYCYANLFTLNKLRESKGMWTKFRPHAGEAGDV 405
DB 61 RRPTKHMPTEQWNTVNFNPAFSYVYCYANLFTLNKLRESKGMWTKFRPHAGEAGDV 120
QY 406 HLAATFLLCHNISHGINLRKSPVLYLYLGOIGLAMPSPNSLSFLDYHRNPFPTFF 465
DB 121 HLAATFLLCHNISHGINLRKSPVLYLYLGOIGLAMPSPNSLSFLDYHRNPFPTFF 180
QY 466 GLNVLSTDDPLQIHLTKPELVEEYSTAASLWKLSSCDLCEIARNYSYQSGFHALKAHW 525
DB 181 GLNVLSTDDPLQIHLTKPELVEEYSTAASLWKLSSCDLCEIARNYSYQSGFHALKSHW 240
QY 526 IGRNYFKRGAGNDIHTNTVPHIRVQPREMIWRNEMKLVYSDNEILLPDEL 576
DB 241 IGRDYKRGDGDNDIHTNTVPHIRVQPREMIWRNEMKLVYSDNEILLPDEL 290
RESULT 14
ID AAG70751
XX AAG70751 standard; Protein; 888 AA.
XX AAG70751;

```

```

XX 27-JUL-2001 (first entry)
XX S cerevisiae apoptosis associated protein YUL035C.
XX Yeast; fungus; apoptosis; infection; proliferative disease;
KW vaccine; autoimmune disease; ischaemia; neurodegeneration.
XX Saccharomyces cerevisiae.
XX WO200102550-A2.
XX 11-JAN-2001.
XX 03-JUL-2000; 2000WO-BE00077.
XX 01-JUL-1999; 99EP-0870141.
XX (JANC ) JANSSEN PHARM NV.
XX Contrexas RH, De Backer MD, Luyten WHML, Malcorps IKL;
PI Nellissen BJM, Reekmans RJ;
XX WPI; 2001-367042/38.
XX N-PSDB; AAH29787.
XX Yeast and fungal nucleic acids encoding proteins involved in a pathway
PT leading to programmed cell death, useful for treating proliferative
PT disorders, yeast and fungal infections, or for preventing apoptosis in
PT certain diseases -
XX Claim 1; Fig 1; 218pp; English.
XX The present invention provides the protein and coding sequences of a
CC number of apoptosis associated proteins from the yeast Saccharomyces
CC cerevisiae and the fungus Candida albicans. These can be used to identify
CC treatments for fungal and yeast infections, for proliferative diseases
CC and for apoptosis related diseases such as autoimmune diseases, ischaemia
CC and neurodegeneration. The present sequence is one of the S. cerevisiae
CC proteins of the invention.
XX SQ Sequence 888 AA;
Query Match 20.2%; Score 627; DB 22; Length 888;
Best Local Similarity 30.1%; Pred. No. 8.2e-53;
Matches 184; Conservative 94; Mismatches 207; Indels 126; Gaps 21;
QY 66 PFTDLHYLRVTAAGNTRTVCHNRLNLLHKKFKHMLNADREFLAQKTAHPHDFNVVRK 125
DB 281 FRDDPAYIELIQSHKFNESVRKLSYLLDKPELFQYLNKSKKEILANKNVYPYRDFVNSRK 340
QY 126 VDTVHHSACMOKHLLRIFIKSKRKEPDEVIFRD---GTWYTLKEVP----- 171
DB 341 VDRDLSLSGCISQRLSEYIWEKINLEPERIV-YQDPETSRKLSURDIFQCGSSNDQPI 399
QY 172 -----ESDLATGYDLNVDLLDVHADKSTFHRFDKENKYNPCGQSRRLREIFLKQDN 222
DB 400 ATGLKLIDDEFID---WYENIYLDIYHLTPNKVAKLVGKEMF-----YLLAKVFLEFDN 451
QY 223 LIQGRFLAEL-TQVFSFDSLASKYQMAEYRIS--IYGRKQSEWDQLASIVNNELHSGNV 279
DB 452 FIEGEYLABIFIKYVHILEKSKYQLAQVSNVQFYSSGEDWYKFSOWLLRWKLVSINI 511
QY 280 VMLVQIPRLY-NVYKEMGIVTSFQNLNDNIFVLPF-----EVTIDPASHPQ---LHVELK 330
DB 512 RWNQIARIFPKLFKE-NVVSFQBFDFIDFNPFLFLEKEQFIDSSVNDIIGLOFFLS 570
QY 331 QVVGLDLVDDESKPE--RRPTKHMPTPEQWT-NVENPAPFSYVYCYANLFTLNKLRESK 387
DB 571 NVCSMDLVIKESDEYYWKEFTDMNCKPKEFTWAGDNPTVAHYMYIYKSLAKVNFRLSQN 630
QY 388 GMTTIFRPHAGE-----AGDVHDLAATFLLCHNISHGINLRKSP----- 427

```







Result No.	Score	Query Match	Length	DB	ID	Description
1	156.5	5.0	516	2	US-09-019-201A-3	Sequence 3, Appli
2	113	3.6	352	4	US-09-328-353-6481	Sequence 6481, Ap
3	111.5	3.6	321	4	US-09-252-991A-28467	Sequence 28467, A
4	106.5	3.4	999	3	US-09-101-8868-2	Sequence 2, Appli
5	105	3.4	464	1	US-07-991-867B-2	Sequence 2, Appli
6	105	3.4	464	1	US-08-107-755A-2	Sequence 2, Appli
7	105	3.4	464	1	US-08-544-332-2	Sequence 2, Appli
8	105	3.4	464	4	US-09-370-861A-2	Sequence 2, Appli
9	104.5	3.4	511	2	US-09-019-201A-2	Sequence 2, Appli
10	100	3.2	604	3	US-09-586-935-3	Sequence 3, Appli
11	100	3.2	604	4	US-09-872-861-4	Sequence 4, Appli
12	99.5	3.2	371	1	US-08-487-748A-12	Sequence 12, Appl
13	99.5	3.2	371	3	US-08-480-070C-12	Sequence 12, Appl
14	99.5	3.2	371	3	US-08-829-525-12	Sequence 12, Appl
15	99.5	3.2	371	3	US-08-609-583A-12	Sequence 12, Appl
16	99.5	3.2	371	3	US-08-937-399-12	Sequence 12, Appl
17	99.5	3.2	371	4	US-09-310-367-12	Sequence 12, Appl
18	99.5	3.2	371	4	US-09-032-337-12	Sequence 12, Appl
19	99.5	3.2	371	4	US-09-464-231-12	Sequence 12, Appl
20	99.5	3.2	523	4	US-09-555-889A-2	Sequence 2, Appli
21	99	3.2	1651	1	US-08-447-411-2	Sequence 2, Appli
22	98	3.2	921	1	US-07-872-644-39	Sequence 39, Appl
23	98	3.2	921	1	US-08-297-494-39	Sequence 39, Appl
24	98	3.2	921	1	US-08-297-510-39	Sequence 39, Appl
25	98	3.2	921	1	US-08-479-532-39	Sequence 39, Appl
26	98	3.2	921	1	US-08-455-526-39	Sequence 39, Appl
27	98	3.2	921	1	US-08-455-526-39	Sequence 39, Appl

Qy	81	NRTVTVCHNRLNLLEHKFKFHLMLNADREFLA---OKTAPH-----RDFVN	122
Dd	143	NL-WVCQGK-----GDKEVIGMRFSKTkPdvATQAdCTWELLSKVRELHG	186
Qy	123	VrkYDTHVHSHSACMNQHLLRFIKSKLRKEPDDEVIFRDGTWTMLKE--VPESLDLTGVD	180
Dd	187	ADKVVDTYL-----REHLTUYPTVKFLDNNEAWEQF--GSIFALLDGGLLFAPSWADYY	237
Qy	181	LNVDLLDVHADKSTPHRRFDKFNLKYNPCQCQSRLREIFLKQDMNLIQGRFLAEI-LTKQVFSD	239
Dd	238	YNA-LKEFHADGVQYLEF-----RSLTPILY-----DLGSTFTELDTVRIYE	280
Qy	240	LSASKYMAEVRISTISYGRKQSEWDOLASWIVNNELHSGNVWLVOIPRLYINVKEMGITV	299
Dd	281	-TLDKY-MAEH-IDFIGSK-----LiYAP-IRNTDKEG----	309
Qy	300	SFOQLLDNI FVPLPEPVITIDPASHPOLHVFLKOVVGILDVDESKPERRPtKMPTEQMT	359
Dd	310	-----LDNYIKVCVEI-----KEKYPDF-----VAGFDLVQGEK--GRELPKDP-	346
Qy	360	NVFNPAFSYAYCYANLFTLNKLRESKM-TTIKFRPHAGBAG-----DYVDHLAA TEL	412
Dd	347	-----IPOLLGMPENIDFYFHAGETNNFGSVTDENLIDA VILL	383
Qy	413	LCHNISHGINLRKSPVLOYLYLGQIGLAMSPLSNNSLFL--DYHRNPPTTFQRGLNVS	470
Dd	384	GTRKRIGHGFALVKGFLVLQMLKERNIAIEVNPISNQVLQLVADYRNHPcAYFPADNPYV	443
Qy	471	LSTDODPLQIHILTKEPLBEYSIAASLWKLSGCCD---LCEIARNVSVQSGFSHALK----	523
Dd	444	ISSDDPSFWKAT--PLSHDFYIAFLGISAHSDMRLLKKLALNSINYSSLSPeQKREVALA	501
Qy	524	HW 525	
Dd	502	KW 503	

## RESULT 2

```

US-09-328-352-6481
; Sequence 6481, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 6481
; LENGTH: 352
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-6481

```

Query Match	3.6%;	Score 113;	DB 4;	Length 352;
Best Local Similarity	18.4%;	Pred. No. 0.0024;		
Matches	84;	Conservative 75;	Mismatches 181;	Indels 116; Gaps 17;
Qy	80	GNTETVCHNRLNLELHKFKFHLMLNADREFLAQKTABHRDFYNNVKYDTHVHHSCAMNQK	139	
Db	2	GNYRVC-----LPHPFFFCGVMNQVELIQ-TLP-----KAEHLVHIISGTEPE	46	
Qy	140	HLRFIKSILRKEPDEVIFRDGYMTLKEVFESLDLTGYDLNLDLVDVHADKSTFFRRFD	199	
Db	47	LMFAIAQRNQIQIP-----YKSVEEVKQAYNFHNLQSFLDIYYAGAN-VLVHEQD	95	
Qy	200	KFNLK---YNPCGSRRLREIFLKDQNLIQGRFLAELTKQVPSDLSASKYQVAEIRSIYG	256	
Db	96	FYDLAWAYFEKCADRV-----VHTEMFDFQTHTRDGIATVING	137	
Qy	257	RKQSEWDQLASWIYNNELHSGNVVLVQIPLRYNNYKEMGIVTISFQNLDDNIFVLPFEVT	316	

[illegible]

### RESULT 3

```

US-09-252-991A-28467
; Sequence 28467, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28467
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28467

```

Query Match	3.6%;	Score 111.5;	DB 4;	Length 321;
Best Local Similarity	26.5%;	Pred. No. 0.003;		
Matches	49;	Conservative	26;	Mismatches 67; Indels 43; Gaps 9
Qy	296	GIVTSFQNLDNITFVPLFEVTIDPASHPOLHVLVQVVGVLVDDESKSPRRPTKHMPTP	355	
Db	137	GLILSP---LRHUSEEQAKTLDQAL--PFRDAFI--AVGLD-----SSEVGH	P 179	
Qy	356	EQWNTVNPAFSYYAYCYANLFTLNKLRSKGWTTIKFRPHAGEAGVDHL--AATFLL	413	
Db	180	SKTQRFVDR-----RSEGLTV---AHAGEGPPYIWEALDLK	217	
Qy	414	CHNISHGINLRKSPVLQVLYVLGQIGLAMSPLSNNSL--FUDYHRNPPPTFFQRCGLNVSL	471	
Db	218	VERIDHGVRAFEDERLMRLRILDEQIPLTVCPLSNTKLCVDFDMSQHTLDMLRGVKVTV	277	
Qy	472	STDDP	476	
Db	278	NSDDP	282	

## RESULT 4

```

US-09-101-886B-2
; Sequence 2, Application US/09101886B
; Patent No. 6197507
;
; GENERAL INFORMATION:
;
; APPLICANT: BERG, THOMAS
; APPLICANT: TOLLERSRUD, OLE K
; APPLICANT: NILSSEN, OIVIND
; TITLE OF INVENTION: GENETIC TEST FOR ALPHA-MANNOSIDOSIS
; NUMBER OF SEQUENCES: 104

```

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: BARBARA G. ERNST  
;; STREET: 555 13TH STREET, NW SUITE 701E  
;; CITY: WASHINGTON  
;; STATE: DC  
;; COUNTRY: USA  
;; ZIP: 20004  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent In Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/101,886B  
;; FILING DATE: 29-JANUARY-1998  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION NUMBER: PCT/CB97/00109  
;; FILING DATE: 12-JAN-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: ERNST, BARBARA G  
;; REGISTRATION NUMBER: 30,377  
;; REFERENCE/DOCKET NUMBER: 1181-240  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-783-6040  
;; TELEFAX: 202-783-6031  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 999 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: not relevant  
;; TOPOLOGY: not relevant  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: not relevant  
;; ANTI-SENSE: not relevant  
;;  
US-09-101-886B-2

Query Match 3.4%; Score 106.5; DB 3; Length 999;  
Best Local Similarity 19.7%; Pred. No. 0.073;  
Matches 109; Conservative 70; Mismatches 168; Indels 205; Gaps 28;

QY 9 EKEVINDPCTKPNPNPFT-VYPEPKSEH--VFQTVGVHIVYADKCTESIYPVADATT 65  
DB 276 DKPVVEDTRSPYNKELVRYELKATDQGLYRTKHTVMTMGSD-----FQYENANT 328  
QY 66 FFDLHYLRVTAAGNTRTVCHNRLNLEHFKFPHMLNADREFLAQKTAPHRDFYNVRK 125  
DB 329 WFKNLDKLIQ-----LVNAQR-----ANGIR 350  
QY 126 VDTHVHGSAC-----MNQKHLLRFIKSKURKEPDEVIFRDGYMTLKEVFESLDLTGYDL 181  
DB 351 VNVLYSTACYLWELNKANLSVKK-----DDFFPVADGPF-----WTGY-- 393  
QY 182 NVDDLVDHAKSTFRFDFKFNKYN-----PCGQ-----SRLREIF-- 217  
DB 394 -----FSSRPALKRYER--LSYNFLQVCNLEALAGPAANVPGYSGDSAPLNEAMAV 444  
QY 218 LKQDNLIOQ-----REFLAE-----LTQVFSDLSASKYQMA---EYRISYVG 256  
DB 445 LQHDVAVSGTSEQHVANDYARQLSGWRPCEVILMSNALHLSGLKEDAFCKRLNISICP 504  
QY 257 RQSEWDQLASWVNNELHSGNVVWLQIPRLYNY--KEMGIVTSFQNLNDNIPVPLFE 314  
DB 505 LTQTA--ERFQVIVYNPL-GRKVDWVRLPVSKHVLVKDPG-----GKIVPSDV 551  
QY 315 VTIDPASHPOLHVF--LQKVQGLDLDVDESKEPRTKMTPEQWTVNFPAPFSYAYY 372  
DB 552 VTI-PSSDSQELFSAVPAGVGFSTYSVSQMPNQRPK-----SWSR----- 592  
QY 373 CYANFLTINKLRSGKMTIKFRPHAGEAGDQVDHLAATFL--CHNISHGINLRK 425  
DB 593 ---DLVIONEYLR-----RFDPTGLMLENLEQNLLLPVROAFYWNASTGNLSS 643

QY 426 SPVLQYLYLQIGLAMPSPNNSFLDYHRNPPTFFQRLNVLNLSLSTDDPLQIHLTKPEP 485  
DB 644 QASGAYIF-----RPNQNKPLFVSHW-----AQTHLVKAS 673  
QY 486 LVEEYSIAASLW 497  
DB 674 LVQEVHQNFSAW 685

RESULT 5  
US-07-991-867B-2  
; Sequence 2, Application US/07991867B  
; Patent No. 5476781  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.  
; APPLICANT: Gruidl, Michael E.  
; TITLE OF INVENTION: No. 5476781el Entomopoxvirus Expression System  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/991,867B  
; FILING DATE: 12-DEC-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/14818  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,685  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/657,584  
; FILING DATE: 19-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UFI14.C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 464 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-991-867B-2

Query Match 3.4%; Score 105; DB 1; Length 464;  
Best Local Similarity 22.0%; Pred. No. 0.028;  
Matches 72; Conservative 50; Mismatches 123; Indels 82; Gaps 17;

QY 25 PPTVPEPKS---EHVQTVGVHIVYADKCTESIYPVADATTFFDHLHILRVTAAGN 81  
DB 167 PLGNPKPKAYPRKDKST-----WLSGGDIYNCIYELTMTINTDYDFHLILFEKTDKN 220  
QY 82 TRTVCHN-RNLNLEHFKFPHMLNADREFLAQKTAPHRDFYNVRKVDTHVHHSACMNQKH 140  
DB 221 IATVASSMRCYKLEDRVKFPLMNDKKFFM-----FPIYN-----DFTCCVCDKH 267  
QY 141 LLR-----FTKSKLRKEPDEVIFRDGYMTLK--EVFESLDLTGYD-LNVVDL 185

Db 268 FDKKAAAYFFNSSGIPBELIKQNKYMFIESD---MTIKSHKYNSTNTNVAIYIDV 324  
QY 186 LDVHAD---KSTFHRFDKFNLYN--PCQSRLREIFLKQDNLIQGRFLAELTKOVFS- 238  
Db 325 LSEYLNDFKNNVYFFNTFELQYDSPDCGMFNIIFLYIVVFNKSKF--EFKLYYSM 382  
QY 239 ----DLASAKYQMA-----EYR--ISYGRKQSEWDQOLASWVNNELHSGNVV 280  
Db 383 SFIGDLLASRGALFISRYDINSIDFKNLTLEIFNKKFMELIDMYKKS-----435  
QY 281 WLVPQIPRLYNYKEMGIVTSFQNLNLDN 307  
Db 436 -----NRIMNVCSK--IKNDYDSYIDN 455

## RESULT 6

US-08-107-755A-2  
; Sequence 2, Application US/08107755A  
; Patent No. 5721352  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.  
; APPLICANT: Gruidl, Michael E.  
; TITLE OF INVENTION: No. 5721352el Entomopoxvirus Expression System  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: Florida  
; COUNTRY: U.S.A.  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/107,755A  
; FILING DATE: 19-AUG-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US 07/827,658  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION NUMBER: US 07/657,584  
; FILING DATE: 19-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: UFI114.C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (904) 375-8100  
; TELEFAX: (904) 372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 464 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-107-755A-2

Query Match 3.4%; Score 105; DB 1; Length 464;  
Best Local Similarity 22.0%; Pred. No. 0.028;  
Matches 72; Conservative 50; Mismatches 123; Indels 82; Gaps 17;  
QY 25 PFTYVPEPKS---EHVFQTVGVHIVYADKCTESIYPVADATFTFDLHYILRVTAAGN 81  
Db 167 PLGFNPKPXPYKDKST-----WLSGDIYNCIYPLTMTINDYDYFHLILFEKTDKN 220  
QY 82 TRFVCHN-RUNLLEHKFKHMLNADREFLAQKTAHRDFYNNRVKVDTHVHHSACNNQGH 140  
Db 221 IATVASSMRCYKLEDRVKFLMNDKKRFM-----FPIIYN-----DHFTCCVIDKH 267

QY 141 LLR-----FIKSLRKBPDEVVIFRDGTMYTLK--EVPESLDLTGYD-LNVDL 185  
Db 268 FDKKAAAYFFNSSGIPBELIKQNKYMFIESD---MTIKSHKYNSTNTNVAIYIDV 324  
QY 186 LDVHAD---KSTFHRFDKFNLYN--PCQSRLREIFLKQDNLIQGRFLAELTKOVFS- 238  
Db 325 LSEYLNDFKNNVYFFNTFELQYDSPDCGMFNIIFLYIVVFNKSKF--EFKLYYSM 382  
QY 239 ----DLASAKYQMA-----EYR--ISYGRKQSEWDQOLASWVNNELHSGNVV 280  
Db 383 SFIGDLLASRGALFISRYDINSIDFKNLTLEIFNKKFMELIDMYKKS-----435  
QY 281 WLVPQIPRLYNYKEMGIVTSFQNLNLDN 307  
Db 436 -----NRIMNVCSK--IKNDYDSYIDN 455

## RESULT 7

US-08-544-332-2  
; Sequence 2, Application US/08544332  
; Patent No. 5935777  
; GENERAL INFORMATION:  
; APPLICANT: Moyer, Richard W.  
; APPLICANT: Hall, Richard L.  
; APPLICANT: Gruidl, Michael E.  
; TITLE OF INVENTION: No. 5935777el Entomopoxvirus Expression System  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Gerard H. Bencen  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/544,332  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/991,867  
; FILING DATE: 07-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/107,755  
; FILING DATE: 19-AUG-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO 92/14818  
; FILING DATE: 12-FEB-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,685  
; FILING DATE: 30-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/657,584  
; FILING DATE: 19-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bencen, Gerard H.  
; REGISTRATION NUMBER: 35,746  
; REFERENCE/DOCKET NUMBER: UFI114.C4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 464 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-544-332-2

```

Query Match      3.4%; Score 105; DB 2; Length 464;
Best Local Similarity 22.0%; Pred. No. 0.028;
Matches 72; Conservative 50; Mismatches 123; Indels 82; Gaps 17;

25 PFTVYPEKs---EHVFTVDGVIHVYADKCTESIYPVADATFFDLDHLVILRVTAAGN 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
167 PLGFNPKPAYPRKHDKST-----WLSGGDIYNCIYPLTWINTDYDYFHLILFEKTDKN 222
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
82 TETVCHN-RLNLELHKFKFHLMLNADRFLAQKTAPHRDFYNVRKVDTHVHHSACMNQKH 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 IATVASSRCYKLEDRVXFFLMDKKRFM-----FFLIYN-----DHTCCVIDKH 267
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
141 LLR-----FIKSURKEPDEVVIFRDGYMTLK--EVPESLDLTGYD-LNVLDL 185
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 FDKDKAAAYFFNSSGYIPELIKQNKYMFIESD--MTIKGHKYNSTPNTNAYLYVIDV 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
186 LDVHAD----KSTFHREDKENLKN--PCGOSRLREIEFLKODNLIQGRFLAELTKQVFS- 238
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
325 LSEYLINDIFKNVNYFFNTFELQYDPCDGMFNIFLYIYVYFNIIKSXF--EFKKLYYSM 382
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
239 ----DLSASKYQMA-----EYR--ISYIGRKQSEWDOLASMINVNELHSGNVV 280
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 SPIGDLASSYRGALFISRYDINSIDEPKNTLEIFNICKKFMELIDMYKNS----- 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
281 WLWQIPRLYNNYKEMGIVTFSQNLNDN 307
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
436 ----NRIMNVCSK--IKNDYDSYIDN 455
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

## RESULT 8

```

US-09-370-861A-2
; Sequence 2, Application US/09370861A
; Patent No. 6410221
; GENERAL INFORMATION:
; APPLICANT: Moyer, Richard W.
; APPLICANT: Hall, Richard L.
; APPLICANT: Gruidl, Michael E.
; TITLE OF INVENTION: No. 6410221el Entomopoxvirus Expression System
; FILE REFERENCE: UF114.C4.D1
; CURRENT APPLICATION NUMBER: US/09/370,861A
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: US 07/991,867
; PRIOR FILING DATE: 1992-12-07
; PRIOR APPLICATION NUMBER: US 08/107,755
; PRIOR FILING DATE: 1993-08-19
; PRIOR APPLICATION NUMBER: WO 92/14818
; PRIOR FILING DATE: 1992-02-12
; PRIOR APPLICATION NUMBER: US 07/827,685
; PRIOR FILING DATE: 1992-01-30
; PRIOR APPLICATION NUMBER: US 07/657,584
; PRIOR FILING DATE: 1991-02-19
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Amsacta moorei entomopoxvirus
US-09-370-861A-2

```

Query Match	3.4%;	Score 105;	DB 4;	Length 464;
Best Local Similarity	22.0%;	Pred. No. 0.028;		
Matches	72;	Conservative 50;	Mismatches 123;	Indels 82; Gaps 17;

  

Qy	25	PFTVPEPKS----	EHVFQTVGVHVVADKCTESY	VPVADATFF	FDLHVIIRVTAAGN	81
		:   :   :	:   :   :	:   :   :	:   :   :	
Db	167	PLGFNPKPAYPRKDKST-----	WLSSGGIYNCIYPLTWINTD	YDFYFLILFEKTDKN		220
		:   :   :	:   :   :	:   :   :	:   :   :	
Qy	82	TRTVCHN-RLNLLHFKFPHLMNADREFLA	QKTA	PHRDFVNVYKVDTHVHHSACM	QKH	140
		:   :   :	:   :   :	:   :   :	:   :   :	
Db	221	IATVASSMRCYKLEDRVKFPLMMDKGRFFN-----	FP	I I Y N-----	DHFTCCV	267
		:   :   :	:   :   :	:   :   :	:   :   :	
Qy	141	LIR-----	FIKSLRKEPDEWIFRDTGTYMTLK-	EVFESLDTGYD-LNVDL		185
		:   :   :	:   :   :	:   :   :	:   :   :	

```

268 FDKKKAAAYFNFSSGYTPELTKQNKYMFIESD---NTIKSHKHYNSTPNTNYAYLYIDV 322
186 LDVHAD---KSTFHREDKENLKNY--PCGOSRLREIFLQDNLIQGRFLAELTKQVFS- 238
325 LSEYLNDFIKNVNYFFNTFELQVSDPCGMFNIFLYIYVYFNKSKF--EFKLYYSM 382
239 -----DLSASKYQMA-----EYR--ISYGRKQSEWQDLASWIVNNELHSGNVV 280
383 SFIGDLLASSYRGALFISRYDINSIDEFKNTLEIFNKNKKFMELIDMYKQNS----- 435
281 WLVIQPLRYNYKEMGIWTSFQNLNDN 307
436 -----NRMNVCSK--IKNDYDSYDN 455

RESULT 9
US-09-019-201A-2
; Sequence 2, Application US/09019201A
; Patent No. 5968780
; GENERAL INFORMATION:
; APPLICANT: FENG, PING
; APPLICANT: SOPPET, DANIEL R.
; APPLICANT: LI, YI
; APPLICANT: DILLON, PATRICK J.
; TITLE OF INVENTION: DENDRITIC CELL-DERIVED GROWTH FACTOR
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVENUE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,201A
; FILING DATE: HEREWITH
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: A. ANDERS BROOKES
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF346
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-019-201A-2

```

	Query Match	3.4%	Score 104.5	DB 2	Length 511
	Best Local Similarity	19.4%	Pred. No. 0.038		
	Matches	78	Conservative 51	Mismatches 142	Indels 131
					Gaps 15
Qy	123	VRKYDTHVHSACMNQKHLIRFIKSKLRKEPDEWIFPDGTYMTLKEVFSFL-DLTGYD-	180		
			::::	:	::::
Db	125	VRNVTYRPHCHICTFTRGIMQRFPAHPTRPSE----	KCSKWILLEDYRKRVQNVTFDD	180	
			::::	:	::::
Qy	181	---LNVDLLDVHADKSTFHRFDKFNLYKNPCGQSRLEIFLKQDNLQGRFAELTQVF	237		
			::::	:	::::
Db	181	SLLRNFTLVQHP-----	EVITYQNQVWSXF-----	FTIF	211
			::::	:	::::
Qy	238	SDLASQYQMAEYRISYGRKQSEWDQLASWIVNNELHSGNVVWLQVLPRLYNYKEMG-	296		
			::::	:	::::
Db	212	FTISGLIHVAPRDXVFRSMO-----	EFYEDNLYMETRALLPYVLSGE	258	
			::::	:	::::

QY 297 -----IVTSFONLNDIEV---PLF-EVTIDPASHPOLHV-----FLK 330  
Db 259 HHDEWSVITYEVAQK-FVETHPEFIGIKIYSDHRSDKDVAVIAESRMANGLIKRPFT 317  
QY 331 QVVGDLVDDESKPERRPTKHMPTPEQWTVNFPAPSYAYCYANLFTLNKLRESKGMT 390  
Db 318 VVAGFDLVGHED-----TGHSRDYKEAL 341  
QY 391 TTKFRP-----HAGEAG-----DVDHLAATFLLCHNISHGINLRKSPVLQVLYLIG 436  
Db 342 MIPAKDGVKLPYFFHAGETDWOQTSIDRNILDALMLNTRIGHGFALSCKHPAVRTYSWK 401  
QY 437 QIGLAMSPLSNLSFL--DYHRNPFPTFFQGLNLSLSTDDP 476  
Db 402 DIPIEVCPISNQVLKLVSLRNHPVATLMATCHGPMWISSDDP 443

## RESULT 10

US-09-586-935-3  
; Sequence 3, Application US/09586935  
; Patent No. 6191267  
; GENERAL INFORMATION:  
; APPLICANT: KONG, HUIMIN  
; APPLICANT: HIGGINS, LAUREN S.  
; APPLICANT: DALTON, MICHAEL  
; APPLICANT: KUCERA, REBECCA B.  
; APPLICANT: SCHILDKRAUT, IRA  
; TITLE OF INVENTION: Cloning And Producing The N.BstnBI Nicking Endonuclease  
; FILE REFERENCE: NEB-178  
; CURRENT APPLICATION NUMBER: US/09/586,935  
; CURRENT FILING DATE: 2000-06-02  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 604  
; TYPE: PRT  
; ORGANISM: Bacillus stearothermophilus  
US-09-586-935-3

Query Match 3.2%; Score 100; DB 3; Length 604;  
Best Local Similarity 19.1%; Pred. No. 0.15;  
Matches 45; Conservative 42; Mismatches 78; Indels 70; Gaps 8;  
QY 85 VCHNRLNLEHKKFHLMLNADREF-----LAQKTAPHDFYNNVRKVDTHVHSAC 135  
Db 393 VYHEKKNVIEEKIKARFIANKNTVFEWLTWNGFIILGNALYKKNFVIDEELQPVTHAA- 451  
QY 136 MNQKHLRFIKSKLRKBPDEVIFRDGTMYTLKEVPESLDLTGYDLNVLDLVDH----- 189  
Db 452 -----GNQPDMEIYED--FIVLGEVTTSGATQFKMSESPVTRHYLNKKK 495  
QY 190 -----ADKSTFHRFDKFNKYNPCGQSRRLREIFLKQDNLIQGRFLA 230  
Db 496 ELEKQGVKEKLYCLFIAPENKNTFEFMYKNIQVQ-----TRIPLSLKQFNML-----LM 547  
QY 231 ELTKQVPSDLSASKYQMAEYRISY-----GRKQSE-----WDQLASWIVNNEL 274  
Db 548 VQKLIIEGRRLSSYDIKNLMVSLYRTTIECERKYQIKAGUEETLNNWVVDKEV 602

## RESULT 11

US-09-872-861-4  
; Sequence 4, Application US/09872861  
; Patent No. 6395523  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Yan  
; APPLICANT: Bresnier, Caroline  
; TITLE OF INVENTION: Engineering Nicking Endonucleases From Type IIs  
; FILE REFERENCE: NEB-188  
; CURRENT APPLICATION NUMBER: US/09/872,861

; CURRENT FILING DATE: 2001-06-01  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 604  
; TYPE: PRT  
; ORGANISM: Bacillus stearothermophilus  
US-09-872-861-4

Query Match 3.2%; Score 100; DB 4; Length 604;  
Best Local Similarity 19.1%; Pred. No. 0.15;  
Matches 45; Conservative 42; Mismatches 78; Indels 70; Gaps 8;  
QY 85 VCHNRLNLEHKKFHLMLNADREF-----LAQKTAPHDFYNNVRKVDTHVHSAC 135  
Db 393 VYHEKKNVIEEKIKARFIANKNTVFEWLTWNGFIILGNALYKKNFVIDEELQPVTHAA- 451  
QY 136 MNQKHLRFIKSKLRKBPDEVIFRDGTMYTLKEVPESLDLTGYDLNVLDLVDH----- 189  
Db 452 -----GNQPDMEIYED--FIVLGEVTTSGATQFKMSESPVTRHYLNKKK 495  
QY 190 -----ADKSTFHRFDKFNKYNPCGQSRRLREIFLKQDNLIQGRFLA 230  
Db 496 ELEKQGVKEKLYCLFIAPENKNTFEFMYKNIQVQ-----TRIPLSLKQFNML-----LM 547  
QY 231 ELTKQVPSDLSASKYQMAEYRISY-----GRKQSE-----WDQLASWIVNNEL 274  
Db 548 VQKLIIEGRRLSSYDIKNLMVSLYRTTIECERKYQIKAGUEETLNNWVVDKEV 602

## RESULT 12

US-08-487-748A-12  
; Sequence 12, Application US/08487748A  
; Patent No. 5721351  
; GENERAL INFORMATION:  
; APPLICANT: Levinson, Douglas A.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/487,748A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7853-023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 371 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-487-748A-12

Query Match 3.2%; Score 99.5; DB 1; Length 371;  
Best Local Similarity 21.8%; Pred. No. 0.075;





```

; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-829-525-12

Query Match      3.2%; Score 99.5; DB 3; Length 371;
Best Local Similarity 21.8%; Pred. No. 0.075;
Matches 96; Conservative 49; Mismatches 149; Indels 147; Gaps 22;

QY 165 MTLK---EVFESLDLTGYDINVDLDDVHAD-----KSTFHRFD-KFNLYK-NPCQSRRL 213
Db 1 MTLTAHLSYFLVLLLAGQGLSDSLTKDAGPRPLELKEVFKLFQIRFNRSYWNPAYETRR 60
QY 214 REIFLKQDNLIOGRFL-----AELTKQVFSLSASKYQMAEYRISYGRKQS--EWDQ 264
Db 61 LSIF--AHNLAQAORLQOEDLGTAEFGTPEFSDLTETEEFG-----QLYQERSPERTEN 112
QY 265 LASWIVNNELHSGNVVWLQVIRLYNVYKEMGIVTSFQ-----LTDNIFVPLF 313
Db 113 MTKKVESN-----TWGESVPTCDWRKAKNIISSVKNQGSCKCCKWAMAADNI-QALW 164
QY 314 EVTIDPASHQLHVFLKQVGLDLDVDESKPERRPTKHMPTPQWNTVNFNPAFSYVAYYC 373
Db 165 RIK-----HQQFVDVSQELLD-----CERCNGCNGGFGVWDAY-- 198
QY 374 YANLFTLNKLRKSGMTTKRPHAGEAGVDVHDLAATFLCHNISHGINLRKSPVLQYLY 433
Db 199 ---LTVLN---NSGLASEKDYFQGDGRKPHRCLAKKY-----KKVAMIQ--- 236
QY 434 YLQIGLAMSPLSNNSLFLDYHRNPPFTFFQRLNVLSTDDPLQIHILTKEPLVEEYSIA 493
Db 237 -----DFTMLSNNEQAIAHY-----LAVHCPITVTINMK-LLOHYQKG 273
QY 494 ASLWKLSSCDLCEIARNVSYQSGF-----SHALK-----AHW 525
Db 274 VIKATPSSCDPRQV-DHSVLLVGFGEKEGMGTGTVLSHRKRHRHSSPYWILKNSWGAHW 332
QY 526 IGKNYFKRGPGAGNDIHRNTP 546
Db 333 GEKGYFRLYRGNNTCGVTKYP 353

```

## RESULT 15

```

US-08-609-583A-12
; Sequence 12, Application US/08609583A
; Patent No. 6204371
; GENERAL INFORMATION:
; APPLICANT: Levinson, Douglas A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF IMMUNE DISORDERS
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,583A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,748
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/398,633
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-048
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 371 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-609-583A-12

Query Match      3.2%; Score 99.5; DB 3; Length 371;
Best Local Similarity 21.8%; Pred. No. 0.075;
Matches 96; Conservative 49; Mismatches 149; Indels 147; Gaps 22;

QY 165 MTLK---EVFESLDLTGYDINVDLDDVHAD-----KSTFHRFD-KFNLYK-NPCQSRRL 213
Db 1 MTLTAHLSYFLVLLLAGQGLSDSLTKDAGPRPLELKEVFKLFQIRFNRSYWNPAYETRR 60
QY 214 REIFLKQDNLIOGRFL-----AELTKQVFSLSASKYQMAEYRISYGRKQS--EWDQ 264
Db 61 LSIF--AHNLAQAORLQOEDLGTAEFGTPEFSDLTETEEFG-----QLYQERSPERTEN 112
QY 265 LASWIVNNELHSGNVVWLQVIRLYNVYKEMGIVTSFQ-----LTDNIFVPLF 313
Db 113 MTKKVESN-----TWGESVPTCDWRKAKNIISSVKNQGSCKCCKWAMAADNI-QALW 164
QY 314 EVTIDPASHQLHVFLKQVGLDLDVDESKPERRPTKHMPTPQWNTVNFNPAFSYVAYYC 373
Db 165 RIK-----HQQFVDVSQELLD-----CERCNGCNGGFGVWDAY-- 198
QY 374 YANLFTLNKLRKSGMTTKRPHAGEAGVDVHDLAATFLCHNISHGINLRKSPVLQYLY 433
Db 199 ---LTVLN---NSGLASEKDYFQGDGRKPHRCLAKKY-----KKVAMIQ--- 236
QY 434 YLQIGLAMSPLSNNSLFLDYHRNPPFTFFQRLNVLSTDDPLQIHILTKEPLVEEYSIA 493
Db 237 -----DFTMLSNNEQAIAHY-----LAVHCPITVTINMK-LLOHYQKG 273
QY 494 ASLWKLSSCDLCEIARNVSYQSGF-----SHALK-----AHW 525
Db 274 VIKATPSSCDPRQV-DHSVLLVGFGEKEGMGTGTVLSHRKRHRHSSPYWILKNSWGAHW 332
QY 526 IGKNYFKRGPGAGNDIHRNTP 546
Db 333 GEKGYFRLYRGNNTCGVTKYP 353

```

Search completed: December 15, 2003, 13:29:13  
Job time : 23 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:28:14 ; Search time 36 Seconds  
(without alignments)  
2986.071 Million cell updates/sec

Title: US-10-019-633-2  
Perfect score: 3100  
Sequence: 1 PRRVAPWEKEVINDPCTPK.....NEMKLVSNEILIPDEL 578

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 684280 seqs, 185983659 residues

Total number of hits satisfying chosen parameters: 684280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1534	49.5	801	12	US-10-205-219-187
2	838	27.0	446	15	Sequence 187, App
3	191.5	6.2	160	15	Sequence 19, Appl
4	184	5.9	65	9	Sequence 5825, Ap
5	156.5	5.0	516	12	Sequence 42950, A
6	147	4.7	358	15	Sequence 12950, A
7	111	3.6	1498	12	Sequence 42950, A
8	107	3.5	342	15	Sequence 10132, A
9	106.5	3.4	1032	11	Sequence 8, Appli
10	106.5	3.4	1032	12	Sequence 12440, A
11	104.5	3.4	333	15	Sequence 192, App
12	104.5	3.4	511	12	Sequence 31, Appl
13	103	3.3	1771	12	Sequence 13109, A
14	103	3.3	1771	12	Sequence 2, Appli
15	102	3.3	654	12	Sequence 4, Appli
					Sequence 11, Appl
					Sequence 1854, Ap

16	102	3.3	1715	12	US-10-021-571-4	Sequence 4, Appli
17	102	3.3	1715	12	US-10-117-229-2	Sequence 2, Appli
18	101.5	3.3	1137	12	US-10-259-165-330	Sequence 330, App
19	101	3.3	796	11	US-09-950-041-41	Sequence 41, Appl
20	99.5	3.2	371	12	US-10-252-131-12	Sequence 12, Appl
21	99.5	3.2	371	15	US-10-004-633-12	Sequence 12, Appl
22	99.5	3.2	622	11	US-09-252-088-9	Sequence 9, Appli
23	99	3.2	466	15	US-10-081-872-198	Sequence 198, App
24	99	3.2	682	12	US-10-032-585-7047	Sequence 7047, Ap
25	99	3.2	871	12	US-10-032-585-7050	Sequence 7050, Ap
26	98.5	3.2	1032	15	US-10-088-567-4	Sequence 4, Appli
27	98.5	3.2	1184	12	US-10-117-229-3	Sequence 3, Appli
28	98.5	3.2	1184	12	US-10-117-229-9	Sequence 9, Appli
29	98	3.2	892	10	US-09-895-913A-198	Sequence 198, App
30	98	3.2	921	10	US-09-883-825-39	Sequence 39, Appl
31	98	3.2	921	14	US-10-094-989-5	Sequence 5, Appli
32	98	3.2	942	10	US-09-883-825-43	Sequence 4, Appli
33	97.5	3.1	2471	14	US-10-116-048-4	Sequence 4, Appli
34	96.5	3.1	1032	11	US-09-954-387B-3	Sequence 3, Appli
35	96.5	3.1	1032	12	US-10-272-502A-9	Sequence 9, Appli
36	96.5	3.1	1032	12	US-10-265-072-8	Sequence 8, Appli
37	96.5	3.1	1987	12	US-10-032-585-7518	Sequence 7518, Ap
38	94.5	3.0	904	11	US-09-950-041-6	Sequence 6, Appli
39	94.5	3.0	904	12	US-10-265-072-2	Sequence 2, Appli
40	94.5	3.0	1032	11	US-09-842-758-34	Sequence 34, Appl
41	94.5	3.0	1073	9	US-09-819-249-2	Sequence 2, Appli
42	94	3.0	1148	14	US-10-108-605-67	Sequence 67, Appl
43	94	3.0	1148	14	US-10-108-605-173	Sequence 173, App
44	94	3.0	1148	14	US-10-108-605-275	Sequence 275, App
45	93.5	3.0	1073	15	US-10-157-031-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1

US-10-205-219-187  
; Sequence 187, Application US/10205219  
; Publication No. US20030138803A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pinnock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018200  
; CURRENT APPLICATION NUMBER: US/10/205,219  
; CURRENT FILING DATE: 2002-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; PRIOR FILING DATE: 2001-07-27  
; NUMBER OF SEQ ID NOS: 197  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 187  
; LENGTH: 801  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
; FEATURE:  
; OTHER INFORMATION: AMP deaminase isoform C  
US-10-205-219-187

Query Match	49.5%	Score 1534;	DB 12;	Length 801;
Best Local Similarity	51.8%	Pred. No. 1.7e+140;		
Matches 288;	Conservative 90;	Mismatches 164;	Indels 14;	Gaps 4;
QY	19	PKPNPNPTTYPEPKS-EHVEQTVDGVHVVADKCTE-----SIYPVADATFTFDLHYI	73	
Db	236	PLPQEDPCYCLDDAPPNLGVLVRMQGVLVFDVNDQTLERQPHSLPYDPLETYIVDMSHI	295	
QY	74	LRVTAAGNTRVCNRLNLLLEKPKFHLMLNADREFLAQKTAHRDFNVKRVKDVTHVS	133	
Db	296	LALITDGTPTKTYCHRRNLNLFLESKFSLHMLNEMSEFKELKSNPHRDFNVKRVKDVTHHAA	355	



APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
FILE REFERENCE: Aeonica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 42950  
LENGTH: 65  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AL096773.6  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.99  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.92  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.9  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3  
OTHER INFORMATION: EST HUMAN HIT: AW342074.1, EVALUATE 6.00e-32  
OTHER INFORMATION: SWISSPROT HIT: P23109, EVALUATE 8.00e-33  
US-09-864-761-42950

Query Match 5.9%; Score 184; DB 9; Length 65;  
Best Local Similarity 56.9%; Pred. No. 2.7e-10;  
Matches 37; Conservative: 7; Mismatches 21; Indels 0; Gaps 0;  
QY 126 VDTTHVHSAACMNQKHLRFKSKLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVDL 185  
DB 1 VDTTHAAACMNQKHLRFKSKLRKEPDEVVIFRDGTMTLKEVFESLDLTGYDLNVDL 185  
QY 186 LDVHA 190  
DB 61 LDVHA 65

RESULT 5

US-10-423-907-3  
Sequence 3, Application US/10423907  
Publication No. US20030186383A1  
GENERAL INFORMATION:  
APPLICANT: Fan et al.  
TITLE OF INVENTION: Dendritic Cell-Derived Growth Factor  
FILE REFERENCE: PF346D1C1  
CURRENT APPLICATION NUMBER: US/10/423,907  
CURRENT FILING DATE: 2003-04-28  
PRIOR APPLICATION NUMBER: 09/360,149  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 09/019,201  
PRIOR FILING DATE: 1998-02-05  
PRIOR APPLICATION NUMBER: US 60/038,829  
PRIOR FILING DATE: 1997-02-06  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent in version 3.2  
SEQ ID NO 3  
LENGTH: 516  
TYPE: PRT  
ORGANISM: Sarcophaga perigrina  
US-10-423-907-3

Query Match 5.0%; Score 156.5; DB 12; Length 516;  
Best Local Similarity 22.3%; Pred. No. 3.2e-06;  
Matches 121; Conservative 65; Mismatches 189; Indels 167; Gaps 30;  
QY 25 PFTVPEPKSEHVFTQVDGV-----IHVADKDCSTESIPVADATFTFFDLHYILRVTAAG 80  
DB 88 PHLFTP---SCHLFEVLDDIKQSPFKYISSMPKGAVALHAHDALCSTD--FLIRLTYRD 142  
QY 81 NTRTVCHNRLNLEHKKFKFLMLNADREFLA---OKTAPH-----RDFYN 122  
DB 143 NL-WVCQK-----GDKEVIGMRFSKTPDVATQADCTWELLSSKVRLEHG 186  
QY 123 VRKVDTHVHSAACMNQKHLRFKSKLRKEPDEVVIFRDGTMTLKE--VFESLDLTGYD 180  
DB 187 ADKVDTYL-----REHLYPTVKFLDNNNEQF--GSIFALLDGLLFYAPSWADYY 237  
QY 181 LNVDLLDVHADKSTFHRPDKFNKYNPCGOSRLREIFLKQDNLIQGRFLAEI-TKQVPSD 239  
DB 238 YNA-LKEFHADGVQVLEF-----RSTLPILY-----DLEGTSTELDTVRIYKE 280  
QY 240 LSASKYQMAEYRISYIGRKQSEWDLASWIVNNELHSGNVVWLVQIPRLVNVYKMGIVT 299  
DB 281 -TLDKY-MAEH-IDFIGSK-----LIYAP-IRNTDKEG-----309  
QY 300 SFQNLNDNIFVPLPEVTIDPASHPOLHVLKQVVGLDLDVDESKPERPRTKHMPTPEQWT 359  
DB 310 -----LDNYIKVCVEI---KEKYPDF-----VAGFLVLVQEEK--GRPLKDF-----346  
QY 360 NVFNPAFSYAYCYANLFTLNKLRKSGM-TTIKFRPHAGEAG-----DVDHLAATFL 412  
DB 347 -----IPQLLGMPENIDFYFHAGETWNFGSTVDENLIDAVLL 383  
QY 413 LCHNISHGINLRKSPVQLYLYLQIGLAMSPLSNNSLFL--DYHRPFPPTFFQRLNVS 470  
DB 384 GTRKIGHGFALVGHPLVLQMLKERNIAIEVNPISNQVLQVLVADYRNHPICAYFFADNPV 443  
QY 471 LSTDDPLQIHLTKPLVEEYSIAASLWKLSSCD---LCEIARNVSGSGFSHALK---A 523  
DB 444 ISSDDPFWKAT--PLSHDFYIAFLGIASHSDRLKKLALNSINYSLSPEQKRVALA 501  
QY 524 HW 525  
DB 502 KW 503

RESULT 6

US-10-156-761-10132

```
; Sequence 10132, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10132
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10132

Query Match      4.7%; Score 147; DB 15; Length 358;
Best Local Similarity 23.6%; Pred. No. 1.5e-05;
Matches 100; Conservative 51; Mismatches 156; Indels 116; Gaps 19;

QY 117 HRDFYNNVRKVDTHVHSAACNQHLLRFKSKLRKEPDEVIFRDGTMTLKEVFESLDL 176
Db 20 HSFIAGLPRAELHVVHVSASP-----RIVSELAARHPDSSV-----PTDPEAL----- 63
QY 177 TGYDLNVLDLVHADKSTHRRPKFNLYNCPGQSLRLEIFLKQDNLIQ-GRFLAELTKQ 235
Db 64 -----ADYFTTDFAHF-----IKVYLSVVDLIRTPEDVRLTYE 98
QY 236 VFDLSASKYQMAEYRISYGRKQSEMDQLASWIV-----NNELHSGNVV-WLVQIPRL 288
Db 99 VARELARQQRYAELTITPFSSTRGIDERAFMDAIEDARKSAEAFGVVLRWCFDIP-- 156
QY 289 YNVYKEMGIVTSFQNLNDNIFVPLFEVTTDPASHQPLHVFVKQVGLDVLVDDESKPERRP 348
Db 157 -----GEAGLESA-----EETVRLAT-----DORLREGLV 182
QY 349 TKHMPTEBQWTVNFPNPAFSYAYVCYANLFTLNKLRSGMTIKRPHRAGEAGVDHL- 407
Db 183 SFGLGGPE--IGVPRPQKPYFDRAIA-----AGLRSV---PHAGETTGPEYVW 226
QY 408 -AATFLLCHNISHGINLRKSPVLQYLYLQ--IGLAMSPLSNLSL-----FLDYHNPFP 460
Db 227 DAUTDLRAERIGHGTSSAQDPKL--LAHLAEHRIPLEVCTPSNIATRAVRLTDEH--PVK 282
QY 461 TFFQRLNLSLTDPLQLHLTKRPLVEEYSIAASLWKLSSCDLCEIARNVYVQSGFSHA 520
Db 283 EFVRAGVTVVINSDDP---PMFGTDLNNEVAIAARLLDLDERGLAGLAKNSVEASFLDAA 339
QY 521 LKA 523
Db 340 GKA 342

RESULT 7
US-10-021-571-8
; Sequence 8, Application US/10021571
; Publication No. US20030166056A1
; GENERAL INFORMATION:
; APPLICANT: CHAO, Moses V.
; APPLICANT: KONG, Haeyoung
; TITLE OF INVENTION: A TRANSMEMBRANE PROTEIN AS A DOWNSTREAM TARGET OF NEUROTROPHIN AN
; TITLE OF INVENTION: RECEPTOR TYROSINE KINASES, DNA ENCODING SAME AND MONOCLONAL ANTI
; FILE REFERENCE: CHAO11A
; CURRENT APPLICATION NUMBER: US/10/021,571
```

```
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,909
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Drosophila
US-10-021-571-8

Query Match      3.6%; Score 111; DB 12; Length 1498;
Best Local Similarity 18.2%; Pred. No. 0.47;
Matches 90; Conservative 75; Mismatches 173; Indels 156; Gaps 21;

QY 26 FTYVPEPKSEHVQTVDOGVHIVYADKCTESIYPVADATTFETDLHYILR----- 75
Db 703 FTFADKEKEHILVALYVIAAVMGTLICTHLHLVLAKEVFSLTSHIRLVKKAARSSAP 762
QY 76 VTAAGNTRTVCHNRLNLE-----HKFKFHLMLNADREFLAQKTAP 116
Db 763 LTMIGAFAVAVMTDMVKCLDAFTNQOSRLGVVIDALDSCDTERILTLNNAVQTL--SSP 819
QY 117 HRDFYNNVRKVDTHV-----HHSACMNQKHLLRFKSK-----KLKREP 153
Db 820 NRPFVLLITSVDPHVIKAAEAANSRRRLFTEGGIGGHDFLRLNLVHLFPVYLQNSGLRKVQRAQ 879
QY 154 DEVIFR---DGTMYT-----LKEVFESLDLT-----GYDLNVLDL 185
Db 880 MTAFLFKSGGGDYQTDGPTLGHVSARRLSNASEIISQEKLRGPARGGGKKLRUSE 939
QY 186 LDVHADKSTHRRPKFNLYNCPGQSLRLEIFLKQDNLI--IQGRFLAELTKQVFS--L 240
Db 940 SVASGTGNLIR-----LGQNPQTVLDLSRVLTDYEDVNPMSRMLMNVIIYITVRL 994
QY 241 SASKYQMAEYRISYGRKQSEMDQLASWIV--NNELHSGNVVWLVQIPRLYNNYK-EMGI 297
Db 995 KAFQIEFSWYRLSSWNLTEQWPLRASMIVLHHQDFMDSNADESVSLSQSVYEKLRPKLAY 1054
QY 298 VTSFQNLID-----NIFVPLFEVTTDPASHQPLHVFVKQV 333
Db 1055 LREAPLLELDRDERKLDALQLHKSLLVADLRIFLP-FTINLDP-----YLRKVL 1105
QY 334 GLD--LVDE-----SKPE-----RRPTKHMPTPEQWTVNFPNPAFSYAYVCYANL 377
Db 1106 KEDQOTIEDGSLVLAQARPSVNTMRQFPATTVYVPSQ-----AYPPYQMFQNE 1155
QY 378 FTLNKLRESKGMT 391
Db 1156 YPANELR-SRNLST 1168

RESULT 8
US-10-156-761-12440
; Sequence 12440, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
```

SEQ ID NO	12440	12441	12442	12443	12444	12445	12446	12447	12448	12449	12450	12451	12452	12453	12454	12455	12456	12457	12458	12459	12460	12461	12462	12463	12464	12465	12466	12467	12468	12469	12470	12471	12472	12473	12474	12475	12476	12477	12478	12479	12480	12481	12482	12483	12484	12485	12486	12487	12488	12489	12490	12491	12492	12493	12494	12495	12496	12497	12498	12499	12500	12501	12502	12503	12504	12505	12506	12507	12508	12509	12510	12511	12512	12513	12514	12515	12516	12517	12518	12519	12520	12521	12522	12523	12524	12525	12526	12527	12528	12529	12530	12531	12532	12533	12534	12535	12536	12537	12538	12539	12540	12541	12542	12543	12544	12545	12546	12547	12548	12549	12550	12551	12552	12553	12554	12555	12556	12557	12558	12559	12560	12561	12562	12563	12564	12565	12566	12567	12568	12569	12570	12571	12572	12573	12574	12575	12576	12577	12578	12579	12580	12581	12582	12583	12584	12585	12586	12587	12588	12589	12590	12591	12592	12593	12594	12595	12596	12597	12598	12599	12600	12601	12602	12603	12604	12605	12606	12607	12608	12609	12610	12611	12612	12613	12614	12615	12616	12617	12618	12619	12620	12621	12622	12623	12624	12625	12626	12627	12628	12629	12630	12631	12632	12633	12634	12635	12636	12637	12638	12639	12640	12641	12642	12643	12644	12645	12646	12647	12648	12649	12650	12651	12652	12653	12654	12655	12656	12657	12658	12659	12660	12661	12662	12663	12664	12665	12666	12667	12668	12669	12670	12671	12672	12673	12674	12675	12676	12677	12678	12679	12680	12681	12682	12683	12684	12685	12686	12687	12688	12689	12690	12691	12692	12693	12694	12695	12696	12697	12698	12699	12700	12701	12702	12703	12704	12705	12706	12707	12708	12709	12710	12711	12712	12713	12714	12715	12716	12717	12718	12719	12720	12721	12722	12723	12724	12725	12726	12727	12728	12729	12730	12731	12732	12733	12734	12735	12736	12737	12738	12739	12740	12741	12742	12743	12744	12745	12746	12747	12748	12749	12750	12751	12752	12753	12754	12755	12756	12757	12758	12759	12760	12761	12762	12763	12764	12765	12766	12767	12768	12769	12770	12771	12772	12773	12774	12775	12776	12777	12778	12779	12780	12781	12782	12783	12784	12785	12786	12787	12788	12789	12790	12791	12792	12793	12794	12795	12796	12797	12798	12799	12800	12801	12802	12803	12804	12805	12806	12807	12808	12809	128
-----------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-----

359	Qy	----	TNVFN--PASYAYAYCYANLFTLNLK--RESKGMTTKIKRPHAGEAGD	403
497	Db	DIACNLNSFNANTQVFNCTEFSMPHIKYLDL--TNRLDFDDNNAFSDL----	HDLEVLDD	551
404	Qy	VDLAATFLULCHNISHGINLRKSPVLQYLYVLGOIGLAMPLSNNLSF-----		451
552	Db	LGNHAHYFSIA--GVTHRLGF-----IQNL-----INLRVLNLSHGVIYTLTESELSKSIS		600
452	Qy	-----LDYHEN-----PPTPFF-----ORGLNVSLSLTDPLQI-----	HLTKSP	485
601	Db	LKELVPSGNRLDHLNANDGKWSIFKSLQNLRLDLSYNNLQOI--PNGAFNLPLQSLQL		660
486	Qy	LVEEYSIAASLKL-----SSCDLCEIARNSVY-----QSGFSHALKAHWIGKNYFKRQGA		536
661	Db	LISGNKLREFFNNTLLQYFPFHLLDLDSRNELYFLPNCLSKFAHSLTLLLSHNHFHSLPS		720
537	Qy	GNDIHRNTNVPHIRVQOFREMIWRNEMKVL	564	
721	Db	GLFSARNLVHLDLSF-----NTIKMI	742	

## RESULT 11

US-10-156-761-13109  
; Sequence 13109, Application US/10156761  
; Publication No. US20030119018A1

```

:
: GENERAL INFORMATION:
: APPLICANT: OMURA, SATOSHI
: APPLICANT: IKEDA, HARUO
: APPLICANT: ISHIKAWA, JUN
: APPLICANT: HORIKAWA, HIROSHI
: APPLICANT: SHIBA, TADAYOSHI
: APPLICANT: SAKAKI, YOSHIYUKI
: APPLICANT: HATTORI, MASAHIRA
: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
: FILE REFERENCE: 249-262
: CURRENT APPLICATION NUMBER: US/10/156,761
: CURRENT FILING DATE: 2002-05-29
: PRIOR APPLICATION NUMBER: JP 2001-204089
: PRIOR FILING DATE: 2001-05-30
: PRIOR APPLICATION NUMBER: JP 2001-272697
: PRIOR FILING DATE: 2001-08-02
: NUMBER OF SEQ ID NOS: 15109
: SEQ ID NO 13109
: LENGTH: 333
: TYPE: PRT
: ORGANISM: Streptomyces avermitilis
: US-10-156-761-13109

```

Query Match 3.4%; Score 104.5; DB 15; Length 333;  
Best Local Similarity 20.9%; Pred. No. 0.2;  
Matches 66; Conservative 36; Mismatches 117; Indels 97; Gaps 12;

Qy	257	RKQSEWOLASWIVNNELHSGNVVLVQIPRLYNVYKEMGIWTSQNLLD--NIFVPL--	312
Db	41	RKAYQFDLOQSL-----NLY--YELMAVLRTQDFEDLANAYPRAA	81
Qy	313	-----PQVTTDPASHPOLHVLKQVVG-----LDLVDDSKPBR	346
Db	82	AQGVRAHAEIIFDPQAHARGVAMGTIVVGLMPLALGRSTENHGVSTQLMCFURDSEASA	141
Qy	347	RPTKHMPPPE-----QWNTVFNPAFSYAYYCYANLFTLANKLRSGMTTIKFRPH	397
Db	142	LETLEAAKPYLDRIIGILDSAEUVGHPAKFREVEYAAAAALGLRRV-----AH	189
Qy	398	AGEAGDVVDHL--AATFLLCNMISHGINLRKSPVLQVLYLYGLOGIAMSPLSNNSL--F	451
Db	190	AGEGPEYIITEALDVLGVERIDHGLRCMEDDALVERLVRVRVPITLCPLSNVRLEAVDV	249
Qy	452	LDYHRNPPTTFQFGLNVSLSTDDPL-----QIHLTKEPVEBYSAAISLWKLSCD	503
Db	250	LAESH--PLPAMLDAGLLCTVNSDDPAYFGFYVADNFHAVRDAL-----GLDQBR	296

Qy	504 LCEARNSVYQGFSH	519
Dd	297 MRELARNSFVASFLEH	312

RESULT 12

```

US-10-423-907-2
;
; Sequence 2, Application US/10423907
; Publication No. US20030186383A1
; GENERAL INFORMATION:
; APPLICANT: Fan et al.
; TITLE OF INVENTION: Dendritic Cell-Derived Growth Factor
; FILE REFERENCE: EP346D1C1
; CURRENT APPLICATION NUMBER: US/10/423,907
; CURRENT FILING DATE: 2003-04-28
; PRIOR APPLICATION NUMBER: 09/360,149
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 09/019,201
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: US 60/038,829
; PRIOR FILING DATE: 1997-02-06
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 511
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-423-907-2

```

Query Match 3.4%; Score 104.5; DB 12; Length 511;  
Best Local Similarity 19.4%; Pred. No. 0.38;  
Matches 78; Conservative 51; Mismatches 142; Indels 131; Gaps 15;

123 VRKVDTHVHSACNMQKHLRPIKSLRKPEDEVIFRDGTMTLKVEPFESL-DLTGYD- 180  
125 VAVVTYRPHCHICFTPRGIMQFRAHPTPRSE-----KCSKWILLEDRKEVQNVTFEDD 180  
181 ---LNVDDLVDHADKSTPHRFDEKFNLYKNPCQSGRLREIFLQKQONLQGRFLABELTKQVF 237  
181 SLRLRFTLVTORP-----EVIYTNQNVVMSKF-----ETIF 211  
238 SDSLASKQMAEYRISYIGRKQSEWDQLASIVNNELHSGNVVWLQIPLRYNLYKEMG- 296  
212 FTISGLIHVAPFRDYVFRSNQ-----EYEDNVLYMEIRARLLPYVLSGE 258  
297 -----IVTSQNLLDNITFV---PLF-EVTTIDPASHFQLHV-----FLK 330  
259 HHDEEWSKYTQEVAAQK-FVTHEPFEGIKIYYSDHRSKDVAVIAESIRWAMGLURIKFPT 317  
331 QVVGGLDLVDDESKPERRPTKHMPTPEQWNTVNFNPAFSYAYCYANLFTLNKLRESKGMT 390  
318 VVAGEDLVGHED-----TGHSRLDYREAL 341  
391 TTKFRP-----HAGEAG-----DVDHLAATFLCHNISHGINLRKSPVLQYLYVLG 436  
342 MIPAKDGVKLPYFFHAGETDQGTSDRNILDALMLNTRIGHGFALSCKHPAVRTYSWK 401  
437 QIGLAMSPLSNNSLFL---DYHRNPPTFFQFGLNVSJLSTDDP 476  
402 DIPIEVCFISNOVLKLVSDLRNHPVATLMTGHPMVISSDDP 443

RESULT 13  
S-10-117-229-4  
Sequence 4, Application US/10117229  
Publication No. US20030190625A1  
GENERAL INFORMATION:  
APPLICANT: Origene Technologies, Inc.  
TITLE OF INVENTION: Human Kidings220PC  
FILE REFERENCE: 9U 104 R1  
CURRENT APPLICATION NUMBER: US/10/117,229  
CURRENT FILING DATE: 2002-04-08  
NUMBER OF SEQ ID NOS: 11



; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 4									
; LENGTH: 1771									
; TYPE: PR									
; ORGANISM: Homo sapiens									
US-10-117-229-4									
Query Match 3.3%; Score 103; DB 12; Length 1771;									
Best Local Similarity 19.9%; Pred. No. 3.7;									
Matches 107; Conservative 61; Mismatches 211; Indels 158; Gaps 25;									
QY	60	VADATFTFDHVLIVRTAAGNTRVCHNRLNLEHKKF----	FHLMNADREFLAQKAP	116					
Db	705	VLCNRTWQVDSLL-----NSQRKLHNAASKL-HKLKSEGFMKVLKCEVELMARM-----		755					
QY	117	HRDFYNVRKVDTHVHHSACMNQKHLRFIKSKLRKEPDEV-----	VIFRDGYTM	165					
Db	756	-----AKTID-----SFTQNTRLWIIDGLDACEQDKVQLMDTVRVLFSGKGFIAIF	804						
QY	166	-----TLKEVPESLDLTGYDL--NVDLLDVHADKSTHFRDKFNLYN-----		206					
Db	805	ASDPHIIIKAINQNLSNVLKRSNNGHDMYRNIVHLPVFLNSRGLSNARKFLVTSATNGD	864						
QY	207	-PC-----GQSRLEI-----FLKQDNLIQGRFLAELTKOV-----		236					
Db	865	VPCSDTTGIOEDADDRVSONSLGEMTKLGSKTALNRDYYRRQWORTITROMSFDLT	924						
QY	237	-----FSDLSASKYQMAEYRISIIYGRK-----QSEWDQLASWIVNNELHSGNVVWL	283						
Db	925	LVTEDWFSDISQPMRRLNIVSVTGRLLRANQISFNMDRLASWINLTQWPYRTSWLIL	984						
QY	284	QIPRLYNNVYKEMGIIVTSFQNLDDNI-----FVPLFEVITDPASHPOLHVLKQVVGDLV	338						
Db	985	YLEETEGIPDQMTLKTIIYERISKNIPTTKDVEPLLEIDGDIRN---FEVFLSSRTPVLVA	1041						
QY	339	DDESKPERRPTKHMPTPEQWTVNFPAPSYAYCYANL-----FTLNKLRSGKMT	390						
Db	1042	RD-----VKVFLP-----CTVNLDPKLREIIADVRAAREQISIG	1075						
QY	391	TIKFRP---HAGEAGDVHDLAATFLLCHNISH-----GINLRKSPVLQYLYLGGQIGLA	441						
Db	1076	GLAYPPLPLHEGPPRAPSGYSQPPSVCSSTSFNGFPAGGV---VSPQPHSSYSGMTG-P	1131						
QY	442	MSPLNNSLFLDYHRNP--FPTFFQR-----GLNVSLSLTDPLQIHLTKPELVEEYS	491						
Db	1132	QHPFYNRPFAPLYTPRYPGSQHLISRPSVKTSLPDQNGNLEVIKEDAAEGLS	1188						
RESULT 15									
US-10-094-749-1854									
; Sequence 1854, Application US/10094749									
; Publication No. US20030219741A1									
; GENERAL INFORMATION:									
; APPLICANT: ISOGAI, TAKAO									
; APPLICANT: SUGIYAMA, TOMOYASU									
; APPLICANT: OTSUKI, TETSUJI									
; APPLICANT: WAKAMATSU, AI									
; APPLICANT: SATO, HIROYUKI									
; APPLICANT: ISHII, SHIZUKO									
; APPLICANT: YAMAMOTO, JUN-ICHI									
; APPLICANT: ISONO, YUUKO									
; APPLICANT: HIO, YURI									
; APPLICANT: OTSUKA, KAORU									
; APPLICANT: NAGAI, KEIICHI									
; APPLICANT: IRIE, RYOTARO									
; APPLICANT: TAMECHIKA, ICHIRO									
; APPLICANT: SEKI, NAOKIKO									
; APPLICANT: YOSHIKAWA, TSUTOMU									
; APPLICANT: OTSUKA, MOTOYUKI									
; APPLICANT: NAGAHARI, KENJI									
; APPLICANT: MASUHO, YASUHIKO									
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA									
; FILE REFERENCE: 08435/0160									
; CURRENT APPLICATION NUMBER: US/10/094,749									
; CURRENT FILING DATE: 2002-03-12									
; PRIOR APPLICATION NUMBER: 60/350,435									
; PRIOR FILING DATE: 2002-01-24									
; PRIOR APPLICATION NUMBER: JP 2001-328381									
; PRIOR FILING DATE: 2001-09-14									
; NUMBER OF SEQ ID NOS: 3381									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 1854									
; LENGTH: 654									
; TYPE: PR									
; ORGANISM: Homo sapiens									
US-10-094-749-1854									

; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 4									
; LENGTH: 1771									
; TYPE: PR									
; ORGANISM: Homo sapiens									
US-10-117-229-4									
Query Match 3.3%; Score 103; DB 12; Length 1771;									
Best Local Similarity 19.9%; Pred. No. 3.7;									
Matches 107; Conservative 61; Mismatches 211; Indels 158; Gaps 25;									
QY	60	VADATFTFDHVLIVRTAAGNTRVCHNRLNLEHKKF----	FHLMNADREFLAQKAP	116					
Db	705	VLCNRTWQVDSLL-----NSQRKLHNAASKL-HKLKSEGFMKVLKCEVELMARM-----		755					
QY	117	HRDFYNVRKVDTHVHHSACMNQKHLRFIKSKLRKEPDEV-----	VIFRDGYTM	165					
Db	756	-----AKTID-----SFTQNTRLWIIDGLDACEQDKVQLMDTVRVLFSGKGFIAIF	804						
QY	166	-----TLKEVPESLDLTGYDL--NVDLLDVHADKSTHFRDKFNLYN-----		206					
Db	805	ASDPHIIIKAINQNLSNVLKRSNNGHDMYRNIVHLPVFLNSRGLSNARKFLVTSATNGD	864						
QY	207	-PC-----GQSRLEI-----FLKQDNLIQGRFLAELTKOV-----		236					
Db	865	VPCSDTTGIOEDADDRVSONSLGEMTKLGSKTALNRDYYRRQWORTITROMSFDLT	924						
QY	237	-----FSDLSASKYQMAEYRISIIYGRK-----QSEWDQLASWIVNNELHSGNVVWL	283						
Db	925	LVTEDWFSDISQPMRRLNIVSVTGRLLRANQISFNMDRLASWINLTQWPYRTSWLIL	984						
QY	284	QIPRLYNNVYKEMGIIVTSFQNLDDNI-----FVPLFEVITDPASHPOLHVLKQVVGDLV	338						
Db	985	YLEETEGIPDQMTLKTIIYERISKNIPTTKDVEPLLEIDGDIRN---FEVFLSSRTPVLVA	1041						
QY	339	DDESKPERRPTKHMPTPEQWTVNFPAPSYAYCYANL-----FTLNKLRSGKMT	390						
Db	1042	RD-----VKVFLP-----CTVNLDPKLREIIADVRAAREQISIG	1075						
QY	391	TIKFRP---HAGEAGDVHDLAATFLLCHNISH-----GINLRKSPVLQYLYLGGQIGLA	441						
Db	1076	GLAYPPLPLHEGPPRAPSGYSQPPSVCSSTSFNGFPAGGV---VSPQPHSSYSGMTG-P	1131						
QY	442	MSPLNNSLFLDYHRNP--FPTFFQR-----GLNVSLSLTDPLQIHLTKPELVEEYS	491						
Db	1132	QHPFYNRPFAPLYTPRYPGSQHLISRPSVKTSLPDQNGNLEVIKEDAAEGLS	1188						
RESULT 14									
US-10-117-229-11									
; Sequence 11, Application US/10117229									
; Publication No. US20030190625A1									
; GENERAL INFORMATION:									
; APPLICANT: Origene Technologies, Inc.									
; TITLE OF INVENTION: Human Kidins220pc									
; FILE REFERENCE: 9U 104 R1									
; CURRENT APPLICATION NUMBER: US/10/117,229									
; CURRENT FILING DATE: 2002-04-08									
; NUMBER OF SEQ ID NOS: 11									
; SOFTWARE: PatentIn version 3.0									
; SEQ ID NO 11									
; LENGTH: 1771									
; TYPE: PR									
; ORGANISM: Homo sapiens									
US-10-117-229-11									
Query Match 3.3%; Score 103; DB 12; Length 1771;									
Best Local Similarity 19.9%; Pred. No. 3.7;									
Matches 107; Conservative 61; Mismatches 211; Indels 158; Gaps 25;									
QY	60	VADATFTFDHVLIVRTAAGNTRVCHNRLNLEHKKF----	FHLMNADREFLAQKAP	116					
Db	705	VLCNRTWQVDSLL-----NSQRKLHNAASKL-HKLKSEGFMKVLKCEVELMARM-----		755					

Search completed: December 15, 2003, 13:33:44  
Job time : 37 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:21:23 ; Search time 17 Seconds  
(without alignments)  
1598.908 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRVRVAPWEKVINDPCTPK.....NEMKLVYSNDNEILIPDELDL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1680.5	54.2	879	1	Q01433. homo sapien
2	1647	53.1	810	1	P15274. saccharomyc
3	1617	52.2	846	1	P50998. schizosacch
4	1534	50.0	767	1	Q01432. homo sapien
5	1534	49.5	765	1	O09178. rattus norv
6	1531	49.4	766	1	O08739. mus musculu
7	1505	48.5	747	1	P10759. rattus norv
8	1497	48.3	747	1	P23109. homo sapien
9	627	20.2	888	1	P40361. saccharomyc
10	625	20.2	797	1	P38150. saccharomyc
11	350	11.3	88	1	Q02356. rattus norv
12	150	4.8	331	1	Q8e8d4. shewanella
13	149	4.8	334	1	Q97ev1. clostridium
14	138.5	4.5	387	1	O86737. streptomyce
15	121.5	3.9	333	1	Q8fh99. escherichia
16	121.5	3.9	339	1	P53984. streptomyce
17	120.5	3.9	352	1	Q9c19r. lactococcus
18	117.5	3.8	333	1	Q8x661. escherichia
19	117.5	3.8	333	1	P22333. escherichia
20	115.5	3.7	3011	1	P87319. schizosacch
21	112	3.6	331	1	Q8d6q8. vibrio vuln
22	112	3.6	1956	1	Q04956. plasmidom
23	111.5	3.6	316	1	Q916y4. pseudomonas
24	109.5	3.5	7073	1	P59641. h replicase
25	109	3.5	343	1	Q91016. streptomyce
26	106.5	3.4	999	1	Q29451. bos taurus
27	106.5	3.4	1032	1	P58682. mus musculu
28	105	3.4	464	1	P29817. amegacta moo
29	105	3.4	510	1	P58780. sus scrofa
30	105	3.4	511	1	Q9nzk5. homo sapien
31	105	3.4	626	1	Q07744. lactococcus
32	105	3.4	626	1	Q09145. lactococcus
33	104.5	3.4	347	1	P53909. saccharomyc

ALIGNMENTS

RESULT 1

ID	AMD2_HUMAN	STANDARD;	PRT;	879 AA.
AC	Q01433; Q14856; Q14857; Q16686; Q16687; Q16688; Q16729; Q9UDX8;			
AC	Q9UDX9; Q9UMU4;			
DT	01-JUL-1993 (Rel. 26, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	AMP deaminase 2 (EC 3.5.4.6) (AMP deaminase isoform L).			
GN	AMPD2.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=93054531; PubMed=1429593;			
RA	Bausch-Jurken M.T., Mahnke-Zizelman D.K., Morisaki T., Sabina R.L.;			
RT	"Molecular cloning of AMP deaminase isoform L. Sequence and bacterial expression of human AMPD2 cDNA.";			
RL	J. Biol. Chem. 267:22407-22413(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RX	MEDLINE=96103174; PubMed=8526848;			
RA	van den Bergh F., Sabina R.L.;			
RT	"Characterization of human AMP deaminase 2 (AMPD2) gene expression reveals alternative transcripts encoding variable N-terminal extensions of isoform L.";			
RL	Biochem. J. 312:401-410(1995).			
RN	[3]			
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RX	MEDLINE=96350465; PubMed=8764830;			
RA	Mahnke-Zizelman D.K., van den Bergh F., Bausch-Jurken M.T., Eddy R.,			
RT	Sait S., Shows T.B., Sabina R.L.;			
RT	"Cloning, sequence and characterization of the human AMPD2 gene: evidence for transcriptional regulation by two closely spaced promoters.";			
RL	Biochim. Biophys. Acta 1308:122-132(1996).			
CC	-!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY METABOLISM.			
CC	-!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).			
CC	-!- PATHWAY: PURINE NUCLEOTIDE CYCLE.			
CC	-!- SUBUNIT: Homotetramer.			
CC	-!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=4;			
CC	Name=Ex1B-2-3;			
CC	Isold=Q01433-1; Sequence=Displayed;			
CC	Name=Ex1A-2-3;			
CC	Isold=Q01433-2; Sequence=VSP_001271, VSP_001272;			
CC	Name=Ex1A-3;			
CC	Isold=Q01433-3; Sequence=VSP_001274;			
CC	Name=Ex1B-3;			
CC	Isold=Q01433-4; Sequence=VSP_001273;			
CC	-!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP			



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M30449; AAA34420.1; -;  
 DR EMBL; Z46659; CAA86620.1; -;  
 DR PIR; S49744; S49744;  
 DR SGD; S0004498; AMD1;  
 DR GO; GO:0003876; F:AMP deaminase activity; IDA;  
 DR GO; GO:0006163; P:purine nucleotide metabolism; IC;  
 DR GO; GO:0006163; P:AMP deaminase activity; IDA;  
 DR InterPro; IPR006650; A/AMP deam. sub.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR InterPro; IPR006329; AMP deaminase.  
 DR Pfam; PF00962; A deaminase; 1;  
 DR TIGRFAMS; TIGR01429; AMP deaminase; 1;  
 DR PROSITE; PS00485; A DEAMINASE; 1;  
 KW Hydrolase; Nucleotide metabolism.  
 FT ACT SITE 422 422 POTENTIAL.  
 FT ACT SITE 631 631 POTENTIAL.  
 FT ACT SITE 707 707 POTENTIAL.  
 FT ACT SITE 708 708 POTENTIAL.  
 FT CONFLICT 568 568 F > C (IN REF. 1).  
 SQ SEQUENCE 810 AA; 93301 MW; 746DCB43B9B45C93 CRC64;  
 Query Match 53.1%; Score 1647; DB 1; Length 810;  
 Best Local Similarity 53.7%; Pred. No. 2.9e-110;  
 Matches 316; Conservative 83; Mismatches 152; Indels 38; Gaps 3;  
 QY 12 VINDCTPKPNPFTYVPEPKSEHVQF-----VDG----- 43  
 DB 211 VQNDQNNKPKGWVYPPPPKPSYNSDTKTVPVTKPDPAEYVDFKCEIPGPDWDF 270  
 QY 44 -----VIHYVADKDCETSYVPVADATFTFDLHYLRVTAAGNTRVCHNRLNLEH 95  
 DB 271 TLNDDSYVHRSKGTD--ELAQIPTLDYDLEKMSISSDGPFAKSFAYRRLOYLEA 328  
 QY 96 KFKFLMLNADREFLAQKTAPHRDFYVNRKVDTHVHSACMKNQKHLRIKSKLRKEPDE 155  
 DB 329 RNLYLLANEYQETSYSKENPHDFYVNRKVDTHVHSACMKNQKHLRIKSKLRHSDKE 388  
 QY 156 VIFRDGTWTKVFPESLDLFGYDLNVLDLVDHAKDSTFRFDKPNLKNPCGQSLRE 215  
 DB 389 KVIKFDGKLLTLDVFRSLHLAGYDLSIDTLDMAHAKDTFRFDKPNLKNPCIGESRLR 448  
 QY 216 IFKQDNLIQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDLASWVNNELH 275  
 DB 449 IFKTNNTYKGTLYADITKQVIFDLENSKYQCEYRISYGRSLDEWDLASWVINDKYI 508  
 QY 276 SGNVVVLVQIPRLYNYKEMGIVTSFQNLNDNI FVPLFEVITDPASHPOLHVLKQVYGL 335  
 DB 509 SHNRVLVQIPRLYDIYKKTGIVQSFQDICKNLFPQLFEVTKNPSQPKLHVLQVIGF 568  
 QY 336 DLVDESKEPRTKHPPEQWTVNFPASYAYCYANLFTLNKLRKSGKMTTKFR 395  
 DB 569 DSVDDSEKVDREPHRYKPKPLWEAPQNPYSYLYLYSNVASLNQWRAKRGFTVLVR 628  
 QY 396 PHAGEAGVDHLAATFLLCHNTSHGINLRKSPVLQVLYLQIGLAMSPLSNNSFLDYH 455  
 DB 629 PHCAGEAGDPEHLVSAVLLAHGSHGILLRKVKFPVQVLYLDQVGLAMSPSLNNAFLTYD 688  
 QY 456 RNPFFTFGRGNVLSLTDDEPQIHLTKPELVEEYSIAASLWKLSCDCLCEARNVYQS 515  
 DB 689 KNPPFRYFKGNVLSLTDDEPQLQFSYTRPELVEEYSVAQIYKLNVDNCEARNVLOS 748  
 QY 516 GFSHALKAHWIGKVPKPGAGNDIHRNTNVPHIRVQFREMIRNEMKLV 564  
 DB 749 GWEAQIKKHGWIKDFDKSGVEGNDVVRVTPVDIRINRYDYDTLSTLELV 797

RESULT 3

AMDM\_SCHPO  
 ID AMDM SCHPO STANDARD; PRT; 846 AA.  
 AC P50998; Q9URV7;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-PEB-2003 (Rel. 43, Last annotation update)  
 DE AMP deaminase (EC 3.5.4.6) (Myoadenylate deaminase).  
 GN ADAL OR SPBC106.04.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RN STRAIN=972;  
 RC Rochet M., Levesque H., Gaillardin C.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RN STRAIN=972;  
 RC MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skellon J., Simmonds M., Squares R., Squares S., Stevens K.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Volkart G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RT "The genome sequence of Schizosaccharomyces pombe."  
 RL Nature 415:871-880(2002).  
 CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
 CC METABOLISM.  
 CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
 CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
 CC -!- SUBUNIT: Homotrimer (By similarity).  
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X91498; CAA62797.1; ALT\_SEQ.  
 DR EMBL; AL110295; CAB53720.1; -;  
 DR PIR; T39261; T39261;  
 DR GeneDB SPombe; SPBC106.04; -;  
 DR InterPro; IPR006650; A/AMP deam sub.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR InterPro; IPR006329; AMP deaminase.  
 DR Pfam; PF00962; A deaminase; 1;  
 DR TIGRFAMS; TIGR01429; AMP deaminase; 1;  
 DR PROSITE; PS00485; A DEAMINASE; 1;

KW Hydrolase; Nucleotide metabolism.  
 FT ACT\_SITE 394 394 POTENTIAL.  
 FT ACT\_SITE 603 603 POTENTIAL.  
 FT ACT\_SITE 679 679 POTENTIAL.  
 FT ACT\_SITE 680 680 POTENTIAL.  
 FT CONFLICT 483 520  
 FT NVRMLQVRLPRLDVYKKGIVTFEVRVNVPEFLFEV ->  
 FT TFVGLFKYLVCMCIRSPVLLRLKRSSEMSLNHCKSF  
 FT (IN REF. 1).  
 SQ SEQUENCE 846 AA; 97454 MW; 701CC540A8B9C25 CRC64;  
 Query Match 52.2%; Score 1617; DB 1; Length 846;  
 Best Local Similarity 57.8%; Pred. No. 4.3e-108;  
 Matches 304; Conservative 84; Mismatches 126; Indels 12; Gaps 3;  
 QY 43 GVTHVYADKDC-----TESYYPVADATFTFDLHYILRVTAAGNTRVCHNRLNLEHKPK 98  
 DB 245 GIPQVYENDSAYIAGTSPFH-IPTIRDYIDLEFLSASSDGSFSPFRLOYLEGRWN 303  
 QY 99 FHLMLNADREFLAQKTAHRDFYVNRKVDTHVHSSACMNQKHLRRIKSLRKEPDEVVI 158  
 DB 304 MYMLLNEYQELADTKKVPKVRDFYVNRKVDTHVHSSALANQKHLRRIKSLRKEPDEVVI 363  
 QY 159 FRDGTWTKVEPESLDLTGYDLNVDLLDVLHAKDSTFHRDPKFNKYNPCGSRRLREIPL 218  
 DB 364 WRDGFELTQEVFDSLSKLTSDYSLDITLDMHAHTDTHFRDFKFNKYNPIGESRLRTIPL 423  
 QY 219 KODNLQGRFLAELTKQVFSLSASKYQMAEYRISYGRKQSEWDOLASWIVNNELHSGN 278  
 DB 424 KTNDINGRYLAELTKVEFTDRTQKYQMAEYRISYGRNREWDKLAAWIDNELFSN 483  
 QY 279 VVWLQIPRLYNYKEMGIVTSFQNLNDIIFVLPFEVITIDPASHPOLHFLKQVGLDLIV 338  
 DB 484 VRWLQIPRLYNYKSGIVEFEEVVRNVFELFEVTKDPRTHPKLHFLQVIGFDSV 543  
 QY 339 DDESKPERPTKMPPEQWNTVNFAPSYAYCYANLFTLNKLRSEKGMTIKFRPHA 398  
 DB 544 DDESKPERPTKMPPEQWNTVNFAPSYAYCYANLFTLNKLRSEKGMTIKFRPHA 603  
 QY 399 GEAGDVHLAATLLCHNLSHGILNRKSPVQVLYLGOIGQAMPSLNSLFLDYHRNP 458  
 DB 604 GEAGDVHLAATLLCHNLSHGILNRKSPVQVLYLGOIGQAMPSLNSLFLDYHRNP 663  
 QY 459 FPTFFORGLNVLSTDDPLQIHILTKPELVEEYSIAASLKLSSCDLCEIARNSVYQSGFS 518  
 DB 664 FLTYFKRGLNVLSTDDPLQFAFTREPLIEEYAVAAQIKLSAVDMCELARNSVLQSGFE 723  
 QY 519 HALKAHWIKNPKFPGAGNDIHRTPVPHIRVOFREMWRNEMKLV 564  
 DB 724 RQLKERWLVDF-----QDIDRTNVPIRLAYRALTLTQETALV 762  
 RESULT 4  
 AMO3\_HUMAN  
 ID AMO3\_HUMAN STANDARD; PRT; 767 AA.  
 AC Q01432;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE AMP deaminase 3 (EC 3.5.4.6) (AMP deaminase isoform E) (Erythrocyte  
 DE AMP deaminase).  
 GN AMPD3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93042002; PubMed=1420359;  
 RA Yamada Y., Goto H., Ogasawara N.;  
 RT "Cloning and nucleotide sequence of the cDNA encoding human  
 RL erythrocyte-specific AMP deaminase.";  
 RL Biochim. Biophys. Acta 1171:125-128 (1992).  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX TISSUE=Keratinocytes;  
 RX MEDLINE=93015995; PubMed=1400401;  
 RA Mahnke-Zizelman D.K., Sabina R.L.;  
 RT "Cloning of human AMP deaminase isoform E cDNAs. Evidence for a third  
 RT AMP gene exhibiting alternatively spliced 5'-exons.";  
 RL J. Biol. Chem. 267:20866-20877 (1992).  
 RN [3]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RX MEDLINE=96201708; PubMed=8611627;  
 RA Mahnke-Zizelman D.K., Eddy R., Shows T.B., Sabina R.L.;  
 RT "Characterization of the human AMPD3 gene reveals that 5' exon usage  
 RT is subject to transcriptional control by three tandem promoters and  
 RT alternative splicing.";  
 RL Biochim. Biophys. Acta 1306:75-92 (1996).  
 RN [4]  
 RP SEQUENCE FROM N.A., AND VARIANT CVS-573.  
 RX MEDLINE=94272475; PubMed=8004104;  
 RA Yamada Y., Goto H., Ogasawara N.;  
 RT "A point mutation responsible for human erythrocyte AMP deaminase  
 RT deficiency.";  
 RL Hum. Mol. Genet. 3:331-334 (1994).  
 CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
 CC METABOLISM.  
 CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
 CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
 CC -!- SUBUNIT: Homotetramer.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=3;  
 CC Name=1B;  
 CC IsoId=Q01432-1; Sequence=Displayed;  
 CC Name=1A;  
 CC IsoId=Q01432-2; Sequence=VSP\_001275, VSP\_001277;  
 CC Name=1C;  
 CC IsoId=Q01432-3; Sequence=VSP\_001276, VSP\_001278;  
 CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP  
 CC DEAMINASE 1 IS THE DOMINANT FORM IN SKELETAL MUSCLE; AMP  
 CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,  
 CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3  
 CC IS FOUND IN ERYTHROCYTES.  
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D12775; BAA02240.1; -  
 CC EMBL; M84720; AAS58365.1; -  
 CC EMBL; M84721; AAS58366.1; -  
 CC EMBL; M84722; AAS58367.1; -  
 CC EMBL; U29926; AAB60410.1; -  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB60410.1; JOINED.  
 CC EMBL; U29910; AAB60410.1; JOINED.  
 CC EMBL; U29911; AAB60410.1; JOINED.  
 CC EMBL; U29916; AAB60410.1; JOINED.  
 CC EMBL; U29917; AAB60410.1; JOINED.  
 CC EMBL; U29918; AAB60410.1; JOINED.  
 CC EMBL; U29922; AAB60410.1; JOINED.  
 CC EMBL; U29924; AAB60410.1; JOINED.  
 CC EMBL; U29925; AAB60410.1; JOINED.  
 CC EMBL; U29926; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29929; AAB60410.1; JOINED.  
 CC EMBL; U29907; AAB60410.1; JOINED.  
 CC EMBL; U29909; AAB



This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U90888; AAC53348.1; ALT\_INIT.  
InterPro: IPR006650; A/AMP deam sub.  
InterPro: IPR001365; A/AMP deaminase.  
InterPro: IPR006329; AMP deaminase.  
Pfam: PF00962; A deaminase; 1.  
TIGRFAMs: TIGR01429; AMP deaminase; 1.  
PROSITE: PS00485; A DEAMINASE; 1.  
Hydrolase; Nucleotide metabolism; Multigene family;  
KW Alternative splicing.

SQ SEQUENCE 765 AA; 88483 MW; F826FA5966221B09 CRC64;

Query Match 49.4%; Score 1534; DB 1; Length 765;  
Best Local Similarity 51.6%; Pred. No. 3.3e-102;  
Matches 288; Conservative 90; Mismatches 164; Indels 14; Gaps 4;

QY 19 PKPNPFTYVPEPKS-EHVFQTVGVHVIYADKDCTE-----SIYFVADATFTFDLHYI 73  
DB 200 PLPQEDPYCLDDAPPNLGVLVRMQGVLVYDQNTMLERQEPHSLPYDPLETVIVDMSHI 259  
QY 74 LRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQKTAHDFYNNVRKVDTHVHHS 133  
DB 260 LALITDGPCTKYCHRLNLFLESFSLHEMLNEMSEFKELKSNPRHDFYNNVRKVDTHHAA 319  
QY 134 ACNQKHLLRFIKSLKRPDEVVIFRDGTMTLKVFESLDITGVDLNVLLDVHADKS 193  
DB 320 ACNQKHLLRFIKYVTEPDRTVAEKLGRKITLRQVDFSLMDPDLDLTVDSLDVHAGQ 379  
QY 194 TFRPDKFNLYKPCQSLREIFLKDQNLIOGRFLAELTKQVFSLSASKYQMAEYRIS 253  
DB 380 TFGDFDKFNKPNVPGASELRLDYLKTENYLGGEYFARMVKEVARELEDSKYQSEPRLS 439  
QY 254 IYGRKQSEWDLASWVNNELHSGNVVWLVIQIRLNNVYKEMGIVTSFQNLNIIIVPLF 313  
DB 440 IYGRSPKWSLAWTFIQHKYVFNMRWIIQVPRYIDFRSKKLLPSFGMLNFIPLF 499  
QY 314 EVTIDPASHPOLHVFVLKVVGLDVLDDSKPERRPTKHM-----PTPEQWTVNFPAPSY 368  
DB 500 QATINQDREHLHLFLKYVTFGFSVDDESK-----HSDHMFSDKSPFDLWTSQNPYPY 555  
QY 369 YAYCYVANFLTKLRESKGMTTKFRPHAGEAGVDVHDLAATFLCHNISHGINKRSPV 428  
DB 556 YLYVMYANIMVNLNRERGLSTFLFRPHCGEAGSITHLVSAFLTADNISHGLLLKSPV 615  
QY 429 LQVLYYLQIGLAMSPLSNNSLFLDYHRNPFPFFQFQRLNLSLSTDDPLQIHLTKPLVE 488  
DB 616 LQVLYYLAQPIAMSPNNSLFLSKPLRFLHKLGLVSLSTDDPMQPHYTKALME 675  
QY 489 EYSIASLWKLSSCDLCEIARNVSGSGFHALKAHWIGKVPKPGAGNDIHRVTNPHI 548  
DB 676 EYAIARQVWKLSTCDLCEIARNVSLQSLGSHQKQKPLGQNYIKEGPEGNDIRKTNVAQI 735  
QY 549 RVQFREMIRWENMKLV 564  
DB 736 RMAFRVETLCLNELSFL 751

RESULT 6

AMD3 MOUSE  
ID AMD3 MOUSE STANDARD; PRT; 766 AA.  
AC O08739; O88692;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE AMP deaminase 3 (EC 3.5.4.6) (AMP deaminase isoform E) (AMP deaminase  
H-type) (Heart-type AMPD).

GN AMPD3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Heart;  
RX MEDLINE=97279027; PubMed=9133604;  
RA Wang X., Morisaki H., Sermuivitayawong K., Mineo I., Toyama K.,  
Ogawara N., Mukai T., Morisaki T.;  
RT "Cloning and expression of cDNA encoding heart-type isoform of AMP  
deaminase.";  
RL Gene 188:285-290 (1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RX MEDLINE=97468276; PubMed=9321472;  
RA Sermuivitayawong K., Wang X., Nagabukuro A., Matsuda Y., Morisaki H.,  
Toyama K., Mukai T., Morisaki T.;  
RT "Genomic organization of Ampd3, heart-type AMPD gene, located in  
mouse chromosome 7.";  
RL Mamm. Genome 8:767-769 (1997).  
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
METABOLISM.  
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
CC -!- SUBUNIT: Homotrimer (By similarity).  
CC -!- TISSUE SPECIFICITY: FOUND IN HEART, LUNG BRAIN, SPLEEN, KIDNEY AND  
TO A LESSER EXTENT IN LIVER.  
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: D85596; BAA19933.1; ALT\_SEQ.  
EMBL: D88994; BAA32548.1; JOINED.  
EMBL: D88985; BAA32548.1; JOINED.  
EMBL: D88986; BAA32548.1; JOINED.  
EMBL: D88987; BAA32548.1; JOINED.  
EMBL: D88988; BAA32548.1; JOINED.  
EMBL: D88989; BAA32548.1; JOINED.  
EMBL: D88990; BAA32548.1; JOINED.  
EMBL: D88991; BAA32548.1; JOINED.  
EMBL: D88992; BAA32548.1; JOINED.  
EMBL: D88993; BAA32548.1; JOINED.  
MGD; MGI:1096344; Ampd3.  
InterPro: IPR006650; A/AMP deam sub.  
InterPro: IPR001365; A/AMP deaminase.  
InterPro: IPR006329; AMP deaminase.  
Pfam: PF00962; A deaminase; 1.  
TIGRFAMs: TIGR01429; AMP deaminase; 1.  
PROSITE: PS00485; A DEAMINASE; 1.  
Hydrolase; Nucleotide metabolism; Multigene family.  
KW CONFLICT 251 251 K -> E (IN REF. 2).  
SQ SEQUENCE 766 AA; 88665 MW; 949B613565EFAAA0 CRC64;

Query Match 49.4%; Score 1531; DB 1; Length 766;  
Best Local Similarity 51.6%; Pred. No. 5.4e-102;  
Matches 287; Conservative 91; Mismatches 164; Indels 14; Gaps 4;

QY 19 PKPNPFTYVPEPKS-EHVFQTVGVHVIYADKDCTE-----SIYFVADATFTFDLHYI 73  
DB 201 PLPQEDPYCLDDAPPNLGVLVRMQGVLVYDQNTMLERQEPHSLPYDPLETVIVDMSHI 260  
QY 74 LRVTAAGNTRTVCHNRLNLEHFKFHLMLNADREFLAQKTAHDFYNNVRKVDTHVHHS 133  
DB 261 LALITDGPCTKYCHRLNLFLESFSLHEMLNEMSEFKELKSNPRHDFYNNVRKVDTHHAA 320



QY 134 ACNMQHLLRFKSKLKEPDEVVIFRDGTYMTLKEVFSLSLDTGVDLNVLDLVDHAKS 193  
 Db 321 ACNMQHLLRFKSKLKEPDEVVIFRDGTYMTLKEVFSLSLDTGVDLNVLDLVDHAKS 380  
 QY 194 TFRPDKFNKYNPCQSLREIFLQKODNLIOGRFLAELTKOVFSLSASKYQMAEYRIS 253  
 Db 381 TFRPDKFNKYNPCQSLREIFLQKODNLIOGRFLAELTKOVFSLSASKYQMAEYRIS 440  
 QY 254 IYGRKQSEWDLASWIVNNELHSGNVVWLQVLPRLVNVYKEMGIVTSFQNLLDNIFVPLF 313  
 Db 441 IYGRSPKSSSLARWFIQHKVYSPNRIIQQVRIYDIIFRSKLLNFCMKLENIFLPLF 500  
 QY 314 EYTIIDPASHPOLHVLKQVVGDLVDDESKPERRPTKM-----PTPEQWTVNFPAPYS 368  
 Db 501 KATINPDHRELHLFLUKYTGTFDSVDDESK---HSDHMFSDKSPSLMTSEQNPYPYS 556  
 QY 369 YAYCYANLFTLNKLRKSGMTTKPRPHAGEAGVDVHLAATFLLCHNLSHGINLRKSPV 428  
 Db 557 YLYMYANLFTLNKLRKSGMTTKPRPHAGEAGVDVHLAATFLLCHNLSHGINLRKSPV 616  
 QY 429 LQVLYLQIGLAMSPLSNLSFLDYHRNPFTFFQGRGLNVSLSTDDPLQIHLTKREPLVE 488  
 Db 617 LQVLYLQIGLAMSPLSNLSFLDYHRNPFTFFQGRGLNVSLSTDDPLQIHLTKREPLVE 676  
 QY 489 EYSIAASLWKLSCDLCEIARNVSQSGFSHALKAHWIGKYNFKRGPAGNDIHRNVPHI 548  
 Db 677 EYAIAAQVWKLSTCDLCEIARNVSQSGFSHALKAHWIGKYNFKRGPAGNDIHRNVPHI 736  
 QY 549 RVQFREMWRNEMKLV 564  
 Db 737 RMAFRYETLCNELSFL 752

## RESULT 7

AMD1 RAT  
 ID AMD1 RAT STANDARD; PRT; 747 AA.  
 AC P10759;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M).  
 GN AMPD1.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 536-548.  
 RC TISSUE=Muscle;  
 RX MEDLINE=87308255; PubMed=3624265;  
 RA Sabina R.L., Marquetant R., Desai N.M., Kalleth K., Holmes E.W.;  
 RT "Cloning and sequence of rat myoadenylate deaminase cDNA. Evidence for tissue-specific and developmental regulation."  
 RT J. Biol. Chem. 262:12397-12400(1987).  
 CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY METABOLISM.  
 CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
 CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
 CC -!- SUBUNIT: Homotetramer.  
 CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE. EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3 IS FOUND IN ERYTHROCYTES.  
 CC -!- SIMILARITY: belongs to the adenosine and AMP deaminases family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).  
 CC -----

DR EMBL; J02811; AAB54086.1; -;  
 DR PIR; A27366; A27366.  
 DR InterPro; IPR006650; A/AMP deam sub.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR InterPro; IPR006329; AMP deaminase.  
 DR Pfam; PF00962; A deaminase; 1.  
 DR TIGRFAMs; TIGR01429; AMP deaminase; 1.  
 DR PROSITE; PS00485; A DEAMINASE; 1.  
 KW Hydrolase; Nucleotide metabolism; Multigene family.  
 FT ACT\_SITE 363 363 POTENTIAL.  
 FT ACT\_SITE 573 573 POTENTIAL.  
 FT ACT\_SITE 649 649 POTENTIAL.  
 FT ACT\_SITE 650 650 POTENTIAL.  
 SQ SEQUENCE 747 AA; 86431 MW; C8928B6F2DD9478 CRC64;  
 Query Match 48.5%; Score 1505; DB 1; Length 747;  
 Best Local Similarity 51.2%; Pred. No. 3.8e-100;  
 Matches 285; Conservative 87; Mismatches 169; Indels 16; Gaps 5;  
 QY 19 PKPNPPTTY--VPEPKSEHVFOITVDGVIHYVADKCTE---STYPVADATTFETDLHY 72  
 Db 188 PKGEDPFRREDLPANLGYHL-KMKGGVYIYIPDEAAASRDEPKPYPNLDDFLDDMNF 246  
 QY 73 ILRTAAGNTVTCNRLNLEHFKFHLMLNADREFLAQKTAHRDFYNNVKVDTHVHH 132  
 Db 247 LLALIAQPVKTYTHRLKFLSSKFQVHOMLNEMDELKELKNPHRDFYNNCRKVDTHIA 306  
 QY 133 SACNQKHLRFIKSKURKEPDEVVIFRDGTYMTLKEVFSLSLDTGVDLNVLDLVDHAK 192  
 Db 307 AACNQKHLRFIKSKVHIDADRVVYSTKEKNLTKELFAQLNMHPYDLTVDSLVDHAGR 366  
 QY 193 STEHFDKFNLYKPCGOSRLREIFLQKODNLIOGRFLAELTKOVFSLSASKYQMAEYRI 252  
 Db 367 QTQRFDFKNDKYNPVGASELDYLTNDYNGEYFATIIKEVGADLVDAKYQHAEPRL 426  
 QY 253 SIYGRKQSEWDLASWIVNNELHSGNVVWLQVLPRLVNVYKEMGIVTSFQNLLDNIFVPL 312  
 Db 427 SIYGRSPDEWKLSSWFGNRIYCPNMTWMLQVRIYDVFRSKNPLPHFGKMLENIFLVP 486  
 QY 313 FEVTIDPASHPOLHVLKQVVGDLVDDESKPERRPTKM-----PTPEQWTVNFPAPYS 367  
 Db 487 FEATINPDHRELHLFLUKYTGTFDSVDDESK---HSGHMFSSKSPKPEETMENNPSY 542  
 QY 368 YVAYCYANLFTLNKLRKSGMTTKPRPHAGEAGVDVHLAATFLLCHNLSHGINLRKSP 427  
 Db 543 YVAYCYANLFTLNKLRKSGMTTKPRPHAGEAGVDVHLAATFLLCHNLSHGINLRKSP 602  
 QY 428 VLQVLYLQIGLAMSPLSNLSFLDYHRNPFTFFQGRGLNVSLSTDDPLQIHLTKREPLV 487  
 Db 603 VLQVLYLQIGLAMSPLSNLSFLDYHRNPFTFFQGRGLNVSLSTDDPLQIHLTKREPLV 662  
 QY 488 EYSIAASLWKLSCDLCEIARNVSQSGFSHALKAHWIGKYNFKRGPAGNDIHRNVPH 547  
 Db 663 EYAIAAQVWKLSTCDLCEIARNVSQSGFSHALKAHWIGKYNFKRGPAGNDIHRNVPH 722  
 QY 548 IRVQFREMWRNEMKLV 564  
 Db 723 IRMAYRYETWCYELNLI 739  
 RESULT 8  
 AMD1 HUMAN  
 ID AMD1 HUMAN STANDARD; PRT; 747 AA.  
 AC P23109;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE AMP deaminase 1 (EC 3.5.4.6) (Myoadenylate deaminase) (AMP deaminase isoform M).  
 GN AMPD1.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCB TaxID=9606;  
 RP [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90264442; PubMed=23451176;  
 RA Sabina R.L., Morisaki T., Clarke P., Eddy R., Shows T.B., Morton C.C.,  
 RA Holmes E.W.;  
 RT "Characterization of the human and rat myoadenylate deaminase  
 RT genes.";  
 RL J. Biol. Chem. 265:9423-9433(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92131279; PubMed=1370861;  
 RA Sabina R.L., Fishbein W.N., Pezeshkpour G., Clarke P.R., Holmes E.W.;  
 RT "Molecular analysis of the myoadenylate deaminase deficiencies.";  
 RL Neurology 42:170-179(1992).  
 RN [3]  
 RP VARIANT LEU-48.  
 RX MEDLINE=92335317; PubMed=1631143;  
 RA Morisaki T., Gross M., Morisaki H., Pongratz D., Zoellner N.,  
 RA Holmes E.W.;  
 RT "Molecular basis of AMP deaminase deficiency in skeletal muscle.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6457-6461(1992).  
 RN [4]  
 RP VARIANTS AMPD DEFICIENCY TRP-388 AND HIS-425.  
 RX MEDLINE=20556569; PubMed=11102975;  
 RA Morisaki H., Higuchi I., Abe M., Osame M., Morisaki T.;  
 RT "First missense mutations (R388W and R425H) of AMPD1 accompanied with  
 RT myopathy found in a Japanese patient.";  
 RL Hum. Mutat. 16:467-472(2000).  
 CC -1- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
 CC METABOLISM.  
 CC -1- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
 CC -1- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
 CC -1- SUBUNIT: Homotetramer.  
 CC -1- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP  
 CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP  
 CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,  
 CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3  
 CC IS FOUND IN ERYTHROCYTES.  
 CC -1- DISEASE: Defects in AMPD1 are the cause of AMP deaminase  
 CC deficiency (AMPD deficiency). It results in exercise-related  
 CC myopathy characterized by muscle aches, cramps, and early fatigue.  
 CC It is one of the most common inherited defects in the Caucasians,  
 CC but not in Asians.  
 CC -1- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M37931; AAG24258.1; -  
 DR EMBL; M37920; AAG24258.1; JOINED.  
 DR EMBL; M37921; AAG24258.1; JOINED.  
 DR EMBL; M37922; AAG24258.1; JOINED.  
 DR EMBL; M37923; AAG24258.1; JOINED.  
 DR EMBL; M37924; AAG24258.1; JOINED.  
 DR EMBL; M37927; AAG24258.1; JOINED.  
 DR EMBL; M37928; AAG24258.1; JOINED.  
 DR EMBL; M37929; AAG24258.1; JOINED.  
 DR EMBL; M37930; AAG24258.1; JOINED.  
 DR EMBL; M60092; AAG24258.1; JOINED.  
 DR PIR; I39444; I39444.  
 DR Genew; HGNC:468; AMPD1.  
 DR MIM; 102770; -  
 DR GO; GO:0003876; F:AMP deaminase activity; TAS.  
 DR GO; GO:0006197; P:adenylate deaminase reaction; TAS.

DR InterPro; IPR006650; A/AMP deam sub.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR InterPro; IPR006329; AMP deaminase.  
 DR Pfam; PF00962; A\_deaminase; 1.  
 DR TIGRfams; TIGR01429; AMP\_deaminase; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; 1.  
 KW Hydrolase; Nucleotide metabolism; Multigene family; Polymorphism;  
 FT ACT\_SITE 363 BY SIMILARITY.  
 FT ACT\_SITE 573 BY SIMILARITY.  
 FT ACT\_SITE 649 BY SIMILARITY.  
 FT ACT\_SITE 650 BY SIMILARITY.  
 FT VARIANT 48 P -> L (POLYMORPHISM; ACTIVITY COMPARABLE  
 FT TO WILD TYPE).  
 FT /FTID=VAR\_013270.  
 FT FT  
 FT VARIANT 388 R -> W (IN AMPD DEFICIENCY; LOSS OF  
 FT ACTIVITY).  
 FT /FTID=VAR\_013271.  
 FT FT  
 FT VARIANT 425 R -> H (IN AMPD DEFICIENCY; LOSS OF  
 FT ACTIVITY).  
 FT /FTID=VAR\_013272.  
 FT FT  
 SQ SEQUENCE 747 AA; 86489 MW; 1E15BEE98B95763 CRC64;  
 Query Match 48.3%; Score 1497; DB 1; Length 747;  
 Best Local Similarity 51.4%; Pred. No. 1.4e-99;  
 Matches 289; Conservative 89; Mismatches 162; Indels 22; Gaps 8;  
 QY 18 TP--KPNENPP--TVPEPKSEHVQTVGVVHVY-----ADKCTESI-YPVADATTF 67  
 DB 185 TTPVKKGDEPPTDNLNENLGLYHL-KMKDGVVYVYVNEAAVSKDEPKPLPVNLD--TFL 241  
 QY 68 TDLHYILRTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQKTAHAPHRDFYVNRKVD 127  
 DB 242 DMNFFLLALIAQGPVKTTHRRLLKFSKFQVHQLNEMDELKELKNPHRDFYVNRKVD 301  
 QY 128 THVHSAACNQKHLRLFIKSLRKEPDEVIFRQDGYMTLKEVPSLDTGYDLNVLLD 187  
 DB 302 THIAAACMNQKHLRLFIKSKYQIDADRVVYSTKEKNLTLELFAKLMHPYDLTVDSL 361  
 QY 188 VHADKSTHTRPDKFNKYNPCQSLREIFLKQDNLIOGRFLAELTKOVPSLSASKYQM 247  
 DB 362 VHAGQTTQRFQKFNKYNPCQSLREIFLKQDNLIOGRFLAELTKOVPSLSASKYQM 421  
 QY 248 AERYISYIGRKQSEWDQLASIVNNELHSGNVNVLVQIPRLYVNYKEMGIVTSFQWLLDN 307  
 DB 422 AEPRLSIYGRSPDESKLSFVFCNRIHCPNWTWIIQVRIYDVFRSKNFLPHFGKMLEN 481  
 QY 308 IFVPLFEVTIDPASHPOLHVLKQVVGDLVDDESKPERRPTKHM-----PTPEQWTVNF 362  
 DB 482 IFMPVFEATINQADPELSPFLKHTGFDSDVDESK---HSGHMFSSKSPKQEWTLK 537  
 QY 363 NPAPSYAYCYANUFLTKLRESKGMTIKRPHAGAGVDVHLLAATFLLCHNLSHGIN 422  
 DB 538 NPSYTYAYYMYANIMVLNLSRKRGMTFLRPHCGAGALTHLMTAFMIADDSLHGLN 597  
 QY 423 LRKSPVLQVLYVGOIGLAMSPLSNLSFLDYHRNPFPTFFORGLNVLSLTDPLQIHLT 482  
 DB 598 LRKSPVLQVLYVGOIGLAMSPLSNLSFLDYHRNPFPTFFORGLNVLSLTDPLQIHLT 657  
 QY 483 KEPLVEEYSIAASLWKLSCDCEIARNVSQSGFSHALKAHWIKGNVFKRGPAGNDIHR 542  
 DB 658 KEPLMEEYAIQAQVFKLSTCDMCEVARNVSLQCGISHEEKVKFLGDNVLEEGPAGNDIRR 717  
 QY 543 TNPVPHIRVQFREMIRNEMKLV 564  
 DB 718 TNVAQIRWAYRYETWCYELNLI 739  
 RESULT 9  
 YJHO\_YEAST  
 ID YJHO\_YEAST PRT; 888 AA.  
 AC P40361;  
 DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 104.3 kDa protein in SMC3-MRPL8 intergenic region.  
 GN YJL070C OR J1095 OR HRD888.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;  
 RX MEDLINE=95282514; PubMed=7762302;  
 RA Vandenbol M., Durand P., Dion C., Portetelle D., Hilger F.;  
 RT "Sequence of a 17.1 kb DNA fragment from chromosome X of  
 Saccharomyces cerevisiae includes the mitochondrial ribosomal protein  
 L8.";  
 RL Yeast 11:57-60 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;  
 RA Sor F.J.;  
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; Z34288; CAA84052.1; -  
 DR EMBL; Z49345; CAA89362.1; -  
 DR EMBL; Z49344; CAA89361.1; -  
 DR EMBL; X88851; CAA61309.1; -  
 DR PIR; S50801; S50801.  
 DR SGD; S0003606; YJL070C.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR Pfam; PF00962; A deaminase; 1.  
 KW Hypothetical protein; Hydrolase.  
 SQ SEQUENCE 888 AA; 104263 MW; DB31A8086224114D CRC64;  
 -----  
 Query Match 20.2%; Score 627; DB 1; Length 888;  
 Best Local Similarity 30.1%; Pred. No. 3.2e-37;  
 Matches 184; Conservative 94; Mismatches 207; Indels 126; Gaps 21;  
 QY 66 FETDLHYILRVTAAGNTRVCHNRLNLEHKKFHLMLNADREFLAQKTAHRDFYVRK 125  
 DB 281 FRDDFAYIIELIQSHKFNVEVRKLSYLLDKPFLFYLNKSKKEILANKVVPYRDFYGRK 340  
 QY 126 VTHVHSAACNMOKHLLRIKSKLRKEPDEVVIFRD---GTMTLKEVF----- 171  
 DB 341 VDRDLSLSCISQORQLSEYIWEKINLEPRIV-YQDPETSKLSLRDIFQFCSSNDQPI 399  
 QY 172 -----ESLDLTGYDLNVDLADVHADKSTFHRFDKFNKYNPCGQSRRLREIFLKQDN 222  
 DB 400 AIGKLIDDEFLD---WYRNILVIDVHLTPNKVAKLVGKEMF-----YLLAKVLEFPDN 451  
 QY 223 LIQGRFLAEL-TKQVFSDSLASAKYQAEYRIS--IYGRKQSEWDQLASHVINNELHSGNV 279  
 DB 452 FIEGYLAEIFIKYVIHILEKSKYQLAQVSNVQFYSSGEDYKFKSQWLLRWKLVSINI 511  
 QY 280 VMLVQIPRLY-NVYKEMGIVTSFQNLNDNIFVPLF-----EVTIDPASHQ---LHVFLK 330  
 DB 512 RNNIARIAPKLFKE-NVVSFQFELDLINPLFTLEKEQIPDISSVNTDIIGQFELS 570  
 QY 331 QVVGDLVDDESKEP---RRPTKXMTPEQWT-NVFNPAFSYIAYCYANLFTLNKLRSEK 387  
 DB 571 NVCSMDLVKESDEYVWKEFTDMCKPKFTWAGQNPVTVAHYMYIYKSLAKVNFRLSQN 630  
 QY 388 GMTTITKFRPHAGE-----AGDVHDLAATFLLCHNISHGINLRKSP----- 427

DB 631 LQNTITLRYCSPSSRTSQFGVDLYFTDQVESLVNCLLCC-----NGGLLQVEPLWDAT 686  
 QY 428 VLQLYLYLGQIGLAMSPLSNNSL-----FL----- 452  
 DB 687 MIQYLYFLQIPILAAPLSSVSLNSQKSTFLKKNKVLLEHDYLDKQDETAKINPSRDITV 746  
 QY 453 ----DYHRNPPTTFPQRLNVSLSTDDPL-QIHLTKEPLVEEYSIAASLWKLSSCDLCRI 507  
 DB 747 GEQRSYETNPFPKMFQMGKISLSKSYLSYNTLEPLIEBSYVAASIIYLLNPTDLCEL 806  
 QY 508 ARNSVYQSGFSHALKAHWIGNK-----YFKRGPAGND-----IHRTNVPHIRVOFR 553  
 DB 807 SRTSVLSGSGYGVYKAHWIGVGVKAPYFEENVGGIDWYDTAKDTSIKHNVPIMRRYR 866  
 QY 554 EMWNRNEMKLV 564  
 DB 867 KETLDQEWNV 877  
 -----  
 RESULT 10  
 YB9Z YEAST  
 ID YB9Z YEAST STANDARD; PRT; 797 AA.  
 AC P38150;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 92.9 kDa protein in SSH1-APE3 intergenic region.  
 GN YBR284W OR YBR2021.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c;  
 RX MEDLINE=94378722; PubMed=8091861;  
 RA Holmstroem K., Brandt T., Kalliesoe T.;  
 RT "The sequence of a 32,420 bp segment located on the right arm of  
 chromosome II from Saccharomyces cerevisiae.";  
 RL Yeast 10:847-862 (1994).  
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X76053; CAA53647.1; -  
 DR EMBL; Z36153; CAA85248.1; -  
 DR PIR; S44546; S44546.  
 DR SGD; S0000488; YBR284W.  
 DR InterPro; IPR001365; A/AMP deaminase.  
 DR Pfam; PF00962; A deaminase; 1.  
 KW Hypothetical protein; Transmembrane; Hydrolase.  
 FT TRANSMEM 627 647 POTENTIAL  
 SQ SEQUENCE 797 AA; 92903 MW; BF854593D0B2C747 CRC64;  
 -----  
 Query Match 20.2%; Score 625; DB 1; Length 797;  
 Best Local Similarity 30.6%; Pred. No. 3.9e-37;  
 Matches 182; Conservative 86; Mismatches 213; Indels 114; Gaps 21;  
 QY 66 FETDLHYILRVTAAGNTRVCHNRLNLEHKKFHLMLNADREFLAQKTAHRDFYVRK 125  
 DB 219 FREDFEWCLKTIRDSLSRSEKSLQVLYNKPFPVQHLHSEEMRQSKVPHKDFYGRK 278  
 QY 126 VTHVHSAACNMOKHLLRIKSKLRKEPDEVV--IFRDGYMTLKEVF-SLDLTG----- 178  
 DB 279 IDLNLISGCSQWLTFEFTWKURKEPDRVIHQAF-NGSHITLSQLFKVNFETFGQFFN 337  
 QY 179 -----YDLNVLDLADVHADKSTFHRFDKFNKYNPCQSRRLREIFL 218

```
Db 338 GLKIIDSFLWKYVIYAKYHLVNDMEIHTGS-----HGKQLRY-----YLIJAKTEL 386
.QY 219 KODNLIQGFLEAELTQ--VPSDLASAKYQMAEYRISI-----YGRQSE--WDQLASWIV 270
Db 387 EFDNYINGBYLAELLTKFLIKPQESKYQLCQLSDVDFQYLYHDNSVDVDMWVFWANWLN 446
.QY 271 NNEHSGNVWLVQIOPRLVNVYKMGIVTSFONLNDNIIVPLF--EVTIDPASHPOLHVF 328
Db 447 HYNIFNNIRWIRISRIYPELYHTKVKVNFQYELNLIKPLFNAENYHLKSLPILKPF 506
.QY 329 LKQWGLDLV--DDESKPERRTKHPTPEQWTN--VFNPAFSYAYCYANLFTLNKLE 385
Db 507 LQSVSSIDLICIDTDNVIWKNFTAVSCLPKDMTSGGDNFTISQMYVYVNVLTNKHIRQ 566
.QY 386 -----SKGWTIIXFRPHAGEAGDVHDLAATFL--CHNISHGINL 423
Db 567 ALHQNTFTLRSSCSPMRTSOFNTLNFTF-----TEAILNNFLLACGFLNAEVL 620
.QY 424 -RKSPVLQYLYLQIGLWPLSN-----NSLFLD-----YHRNPPPTFPQ 464
Db 621 WNAAPSLVLYLQISQPMVAPLNSIVDSKPTMLQEQATGLVLEPSKYKKNPFMKFFE 680
.QY 465 RGLNVLSTDDPLQIH-LTKEPLVEYSIAASLWKLSSCDLCEIARNSVYQSGFHALKA 523
Db 681 MGFKISLSSESILYNNSTKEPIIEYSVAASIYRLHSADLCBELLNSVITSFGSSTLKN 740
.QY 524 HWIG-----KNVFKRGA-----GNDIHTNVPHIRVQPREMIWENMKLV 564
Db 741 KMLGVSLASHDYEVNTGFKWYDCKPNTSLHNVPPIIRQVRSSTLAGEWRLI 795

RESULT 11
AMD2_RAT
ID AMD2_RAT STANDARD; PRT; 88 AA.
AC Q02356;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE AMP deaminase 2 (EC 3.5.4.6) (AMP deaminase isoform L) (Fragment).
GN AMPD2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90307656; PubMed=2365682;
RA Morieaki T., Sabina R.L., Holmes E.W.;
RA "Adenylylate deaminase. A multigene family in humans and rats.";
RL J. Biol. Chem. 265:11482-11486(1990).
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY
CC METABOLISM.
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.
CC -!- SUBUNIT: Homotetramer.
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP
CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP
CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3
CC IS FOUND IN ERYTHROCYTES.
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M38126; AAA40728.1; -.
```

```
DR PIR; A37056; A37056.
DR InterPro; IPR006650; A/AMP deam sub.
DR InterPro; IPR001365; A/AMP deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
KW Hydrolase; Nucleotide metabolism; Multigene family.
FT NON_TER 1
FT ACT_SITE 3 3 POTENTIAL.
FT ACT_SITE 79 79 POTENTIAL.
FT ACT_SITE 80 80 POTENTIAL.
FT NON_TER 88
SQ SEQUENCE 88 AA; 9863 MW; 1AF8DC091DB298E7 CRC64;
Query Match 11.3%; Score 350; DB 1; Length 88;
Best Local Similarity 72.7%; Pred No. 1e-18;
Matches 64; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

.QY 336 PHAGAGDVHDLAATFLCHNISHGINIRKSPVLYLYLQIGLWPLSNNSFLDYH 455
Db 1 PHCGAGPIHLLVSAPFLAENISHGLLRKAPVLYLYLAQIGIAMSPLSNNSFLSYH 60
.QY 456 RNPPTFPFQGLNVSLSTDDPLQIHLYTK 483
Db 61 RNPPEYLSRGLMVSLSTDDPLQHFHTK 88

RESULT 12
ADD_SHEON
ID ADD_SHEON STANDARD; PRT; 331 AA.
AC Q8E8D4;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
GN ADD OR S04731.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MR-1;
RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Uterback T.R., McDonald L.A., C.M.;
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealsen K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF015905; RAN57690.1; -.
```

```
DR TIGR; S04731; -.
```

```
DR HAMAP; MF 00540; -; 1.
```

```
DR Pfam; PF00962; A_deaminase; 1.
```

```
DR TIGRFAMs; TIGR01430; aden_deam; 1.
```

```
DR PROSITE; PS00485; A_DEAMINASE; FALSE NEG.
```

```
KW Hydrolase; Nucleotide metabolism; Complete proteome.
```

```
FT ACT_SITE 197 197 POTENTIAL.
```

[illegible]

```

RESULT 13
ADD_CLOAB
ID_ADD_CLOAB STANDARD; PRT; 334 AA.
AC Q97EVI;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydrolase).
GN ADD OR CAC3005.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OC NCBI_TaxID=1488;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VNM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,
RA Gibson R., Lee H.M., Dubois J., Oiu D., Hitti J., Wolf Y.I.,
RA Tatusev R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
RA Bennett G.N., Koonin E.V., Smith D.R.;
RT "genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum.";
RL J. Bacteriol. 183:4823-4838(2001).
CC -1- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
CC -1- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/submit/).
CC or send an email to license@ebi.ac.uk.
-----
CC

```

```

DR EMBL; AE007797; AAK80946.1; -.
DR PIR; G97269; G97269.
DR HAWAP; MF 00540; -. 1.
DR InterPro; IPR006650; A/AMP_deam_sub.
DR InterPro; IPR001365; A/AMP_deaminase.
DR InterPro; IPR008330; A_deaminase.
DR Pfam; PF00962; A_deaminase; 1.
DR TIGRFAMs; TIGR01430; aden deam; 1.
DR PROSITE; PS00485; A_DEAMINASE; FALSE NEG.
DR Hydrolase; Nucleotide metabolism; Complete proteome.
FT ACT_SITE 200 200 POTENTIAL.
FT ACT_SITE 248 248 POTENTIAL.
FT ACT_SITE 281 281 POTENTIAL.
FT ACT_SITE 282 282 POTENTIAL.
SQ SEQUENCE 334 AA; 38023 MW; FC386C9C8DCD789B CRC64;

Query Match          4.8%; Score 149; DB 1; Length 334;
Best Local Similarity 27.3%; Pred.No. 0.0014;
Matches 65; Conservative % 39; Mismatches 106; Indels -28; Gaps 11;

QY 314 EVTIDPASHQLHVFLKQVV--GLD-LVDDSESKPERRP-----KHMTPEOMTVNFNP 364
DB 100 EIRFAPQHTEQDLNENDVVEALAEALQDGESKLGHNSILCLSRHDPV-ERSIDIUNLV 158
QY 365 AFSYYAYCYANL-----FTLNKLRESKG---TTIKFRPHAGEAGDVHDHLAATFLLC 415
DB 159 ANSYNEGCAVDLAGNESDFPELHKAEADLAYVDNGIKITIHAGETGIAENILKSIIKLH 218
QY 416 --NISHGINURKS-PVLQYLVIYLGIGLAMSPISN--NSLFDYHRNPPTFFPGRLNVLS 470
DB 219 ADRIHGIFAYKSEEILQYVIE-NQVPLEMCPKSNVDTKAVNYKNHPFKKKFYDLGVKVT 277
QY 471 LSTDPPLOIHLTKREPLVEEYSIASLWKLSSCDLCEIARNSVVQSQSHALKAHWICK 528
DB 278 LTNDNR-----VSNSVLVDLYNLNIANTPDFGIIEIKTVIRNGISASFATEEFKNLLKK 332

RESULT 14
ADD1_STRCO
ID_ADD1_STRCO STANDARD; PRT; 387 AA.
AC O86737;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable adenosine deaminase 1 [EC 3.5.4.4] (Adenosine aminohydrolase
DE 1).
GN ADD1 OR SC03662 OR SC6A9.05.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RX MEDLINE=21936410; PubMed=12000953;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2)".
RL Nature 417:141-147(2002).
CC -1- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).
CC -1- SIMILARITY: Belongs to the adenosine and AMP deaminases family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AL031035; CAA19890.1; -;  
 DR PIR; T35436; T35436.  
 DR HAMAP; MF 00540; -; 1.  
 DR InterPro; IPR006650; A/AMP\_deam sub.  
 DR InterPro; IPR001365; A/AMP\_deaminase.  
 DR InterPro; IPR006330; A\_deaminase.  
 DR Pfam; PF00962; A\_deaminase; 1.  
 DR TIGRFAMs; TIGR01430; aden deam; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; FALSE NEG.  
 KW Hydroxylase; Nucleotide metabolism; Complete proteome.  
 FT ACT\_SITE 244 244 POTENTIAL.  
 FT ACT\_SITE 292 292 POTENTIAL.  
 FT ACT\_SITE 325 325 POTENTIAL.  
 FT ACT\_SITE 326 326 POTENTIAL.  
 SQ SEQUENCE 387 AA; 42429 MW; 660F2621C115EA5A CRC64;

Query Match 4.5%; Score 138.5; DB 1; Length 387;  
 Best Local Similarity 22.1%; Pred. No. 0.0099;  
 Matches 95; Conservative 53; Mismatches 157; Indels 125; Gaps 19;  
 QY 108 EFLAQKTAP---HRDFYVNRKVDTHVHSACMNOKHL-----LRFIKSKLRKEPDEWVIFR 160  
 Db 37 EHLVDVDPVDRDLHAFTAGLPKAEHLVHHVGSASPRIVSELAARHADSKVPTDPEALV---- 93  
 QY 161 DGTMTLKEVFESLDLTGYDLNVDLDVHADKSTFHRFDKFLNLYNPGCGSLREIFLQK 220  
 Db 94 --DYFTFTDFAHFDV--YLSVVDLIRTPED----- 120  
 QY 221 DNLQRFALUETKQVFDLSASKYQMAEYRISYGRKQSEWDQLASWVNN-----EL 274  
 Db 121 -----VRLLYEYVARDMARQVRYAELTTPFSSTRGIDEGAFMDAIEDAKAAEA 172  
 QY 275 HSGNVV-WLVQIPRLYVYKEMGIVTSFQNLNDNIFVLFVETIDPASHQLHVFLKQV 333  
 Db 173 EFGTVLRWCFDIP-----GEAGLESA-----ETARLAT----- 201  
 QY 334 GLDLVDDESKPERRPTKHMPTPEQWTVNFPNPAFSYIYCYANLFTLNKLRKSGMTTIK 393  
 Db 202 ----DDRLRPEGLVSFGLGGPE--IGVARPQPKYFDRAIA-----AGLHSV- 242  
 QY 394 FRPHAGEAGDVHL--AATFLCHNISHGINLRKSPVLQVLYLGO--TGLAMSPLSNNS 449  
 Db 243 --PHAGETTPQPTVWEALIDLRAERIGHGTSSAQDEPKL--LAHLAERRIPLEVCPPTSNTA 298  
 QY 450 L-----FLDYHRNPFPPTFFQGLNVLSTDDPLQIHLTKEPLVVEYSIAASLWKLSSCDLC 505  
 Db 299 TRAVRTLDEH--PIKEFVRAGVPTVINSDDP-----PMFGTDLNNEYAAARLLGLDERGLA 353  
 QY 506 ETARNVSYSQS 515  
 Db 354 DLAKNGVEAS 363

RESULT 15  
 ADD\_ECOL6  
 ID ADD\_ECOL6 STANDARD; PRT; 333 AA.  
 AC Q8FH99;  
 DT 15-SEP-2003 (Rel. 42, Created)  
 DT 15-SEP-2003 (Rel. 42, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Adenosine deaminase (EC 3.5.4.4) (Adenosine aminohydroxylase).  
 GN ADD OR C2015.  
 OS Escherichia coli O6.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Escherichia.  
 OX NCBI\_TaxID=217992;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=O6.H1 / CFT073 / ATCC 700928;  
 RX MEDLINE=22388234; PubMed=12471157;  
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,  
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,  
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,  
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;  
 RT "Extensive mosaic structure revealed by the complete genome sequence  
 of uropathogenic Escherichia coli";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).  
 CC -!- CATALYTIC ACTIVITY: Adenosine + H(2)O = inosine + NH(3).  
 CC -!- SIMILARITY: Belongs to the adenosine and AMP deaminases family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE016761; AAN80475.1; -;  
 DR HAMAP; MF 00540; -; 1.  
 DR Pfam; PF00962; A\_deaminase; 1.  
 DR TIGRFAMs; TIGR01430; aden deam; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; 1.  
 KW Hydroxylase; Nucleotide metabolism; Complete proteome.  
 FT ACT\_SITE 197 197 POTENTIAL.  
 FT ACT\_SITE 245 245 POTENTIAL.  
 FT ACT\_SITE 278 278 POTENTIAL.  
 FT ACT\_SITE 279 279 POTENTIAL.  
 SQ SEQUENCE 333 AA; 36360 MW; A89FCBC00F5D3F47 CRC64;

Query Match 3.9%; Score 121.5; DB 1; Length 333;  
 Best Local Similarity 24.7%; Pred. No. 0.13; Mismatches 41; Gaps 7;  
 Matches 47; Conservative 27;  
 QY 318 DPASHPQLHVFL---KQVVGLDLDVDESKPERRPTKHMPTPEQWTVNFPNPAFSYIYCY 374  
 Db 146 EAACQOELEAFLAHRDQITALDLAGDE-----LGFPQSLFLSHF-- 184  
 QY 375 ANLFTLNKLRKSGMTTIKFRPHAGEAGDVHL--AATFLCHNISHGINLRKSPVLQYL 432  
 Db 185 -----NRARDAGWHITV-----HAGEAAGPESIQWAIRELGAERIGHGVKATIEDRALMDF 234  
 QY 433 YVLGIGLAMSPLSN--NSLFLDYHRNPFPFFQGLNVLSTDDPLQIHLTKEPLVVEY 490  
 Db 235 LAEQQIGIESCLTSNIQTSTVADLAHPLKTFLEHGIRASINTDDP----GVQGVDIIEY 291  
 QY 491 SIAASLWKL 500  
 Db 292 TVAAPAAGLS 301

Search completed: December 15, 2003, 13:27:14  
 Job time : 20 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:24:28 ; Search time 40 Seconds  
(without alignments)  
3728.860 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRRVAPWEKVEINDPCTPK.....NEMKLVYSDNEILLIPDEL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL.23.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2644.5	85.3	839	10 O80452	O80452 arabidopsis
2	1798	58.0	743	5 Q9NGX0	Q9ngx0 dictyosteli
3	1687.5	54.4	798	11 Q9DBT5	Q9dbt5 mus musculus
4	1680.5	54.2	798	4 Q9G1A1	Q9g1a1 homo sapien
5	1641	52.9	1008	3 Q9P3N4	Q9p3n4 neurospora
6	1638	52.8	652	5 Q9VY76	Q9vy76 drosophila
7	1638	52.8	707	5 Q91Q07	Q91q07 drosophila
8	1569	50.6	865	5 Q18463	Q18463 caenorhabdi
9	1569	50.6	889	5 Q95ZX5	Q95zx5 caenorhabdi
10	1537	49.6	766	11 Q8CFR4	Q8cfr4 mus musculus
11	1424.5	46.0	481	11 Q922J7	Q922j7 mus musculus
12	1325	42.7	672	5 Q8IE53	Q8ie53 plasmodium
13	1060.5	34.2	1655	5 Q9GRM4	Q9grm4 leishmania
14	954.5	30.8	2093	5 Q9XZY8	Q9xyz8 leishmania
15	901.5	29.1	286	11 Q91Y12	Q91y12 mus musculus
16	581	18.7	288	5 Q9GRR7	Q9grr7 leishmania

17	166	5.4	553	5	Q26642	Q26642 sarcophaga
18	150	4.8	331	16	Q8EBD4	Q8ebd4 shewanella
19	147.5	4.8	543	5	Q8SP8	Q8sp8 dictyosteli
20	141	4.5	364	3	Q8X1T6	Q8x1t6 emericeila
21	138	4.5	339	3	Q9P6J8	Q9p6j8 schizosacch
22	136	4.4	488	5	Q8I3V7	Q8i3v7 plasmodium
23	134	4.3	553	5	Q9VVK5	Q9vvk5 drosophila
24	132.5	4.3	282	2	Q8LOX7	Q8lox7 lactobacill
25	132.5	4.3	502	5	Q95WT8	Q95wt8 culex quinq
26	132	4.3	591	3	Q8NIZ8	Q8niz8 neurospora
27	130	4.2	501	5	Q9VFS0	Q9vfo0 drosophila
28	130	4.2	506	5	Q8WS97	Q8ws97 drosophila
29	121.5	3.9	332	12	Q91GY8	Q91gy8 orf virus.
30	121.5	3.9	333	16	Q8FH99	Q8fh99 escherichia
31	121.5	3.9	334	12	Q9YR3	Q9yrr3 orf virus
32	121.5	3.9	334	12	Q91GY9	Q91gy9 orf virus.
33	120	3.9	4717	3	Q94248	Q94248 schizosacch
34	120	3.9	5890	5	Q8IK84	Q8ik84 plasmodium
35	119.5	3.9	2543	5	Q8IBR8	Q8ibr8 plasmodium
36	118.5	3.8	340	12	Q91GT9	Q91gt9 orf virus.
37	118.5	3.8	340	12	Q91GZ0	Q91gz0 orf virus.
38	118	3.8	506	5	Q9U7C5	Q9u7c5 glossina mo
39	117	3.8	39	6	Q9TUK2	Q9tuk2 sus scrofa
40	117	3.8	442	16	Q8EZR9	Q8ezy9 leptospira
41	116	3.7	1320	5	Q8ILJ3	Q8ilj3 plasmodium
42	116	3.7	3347	5	Q8IEA1	Q8ieal plasmodium
43	116	3.7	4638	5	Q8IK96	Q8ik96 plasmodium
44	115.5	3.7	530	5	Q8T9T6	Q8t9t6 aedes aegypt
45	114.5	3.7	435	16	Q98Q33	Q98q33 mycoplasma

#### ALIGNMENTS

#### RESULT 1

O80452 PRELIMINARY; PRT; 839 AA.  
ID O80452 O80452; Q93ZR9;  
AC AT 01-NOV-1998 (TREMREL. 08, Created)  
DC 01-JUN-2002 (TREMREL. 21, Last sequence update)  
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)  
DE Putative AMP deaminase.  
GN ATG38280 OR F16M14.21/ATG38280.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Magnoliophyta; Streptophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_taxid=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,  
RA Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,  
RA Somerville C.R., Venter J.C.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Lin X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RA Town C.D., Kaul S.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Yamada K., Banh J., Banno F., Dale J.M., Goldsmith A.D., Lee J.M.,  
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Yamamura Y., Yu G.,  
RA Yu S., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,  
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,  
RA Koesema E., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M.,  
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,

RA	Southwick A., Tracy S.E., Shinozaki K., Davis R.W., Ecker J.R.,
RA	Theologis A.;
RT	"Full Length cDNA of gene F16M14.21/At2g38280 (GI:3335375).";
RN	Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RL	[5]
RP	SEQUENCE FROM N.A.
RA	Yamada K., Chang M.M., Chang C.H., Dale J.M., Deng J.M., Hsuan V.W.,
RA	Lee J.M., Quach H.L., Tang C., Toriumi M., Wallender E.K., Wong C.,
RA	Hu H.C., Yu G., Yuan S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA	Kawai J., Kim C., Lam B., Lin J., Miranda M., Narusaka M., Nguyen M.,
RA	Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA	Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT	"Arabidopsis Open Reading Frame (ORF) Clones.";
RL	Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AC093028; AAC27176.2; -
DR	EMBL; AY056301; AAL07150.1; -
DR	EMBL; AV133852; AAM91786.1; -
DR	InterPro; IPR001365; A/AMP_deaminase.
DR	InterPro; IPR006650; A/AMP_deam_sub.
DR	InterPro; IPR006329; AMP_deaminase.
DR	Pfam; PF00962; A_deaminase; 1.
DR	TIGRFAMS; TIGR01429; AMP_deaminase; 1.
DR	PROSITE; PS00485; A_DEAMINASE; 1.
SQ	SEQUENCE 839 AA; _95129 MW; 188F1F4A589A17DA CRC64;
Query Match 85.3%; Score 2644.5; DB 10; Length 839;	
Best Local Similarity 83.9%; Pred.No.5.le-205;	
Matches 480; Conservative 46; Mismatches 45; Indels 1; Gaps 1;	
QY	5 VAPWEKVINDPCTPKENPNPFTVVPKPSEHVQTVDGVTHVVADKDCTESIYPVADAT 64 
Db	267 VAPWEKEIVSDPTPKENTPEFFAHYPOKSDCHCFEMQDGVVHFANKDAKEDLFVADAT 326 
QY	65 TFTFDLHYILRVTAAGNRTRTVCHNRLNLLEHKFEHLMANADREFLAQTKAPHDFYNVR 124 
Db	327 AFTFDLHHVLKVTAAGNIRTLCHRRLLVEQKFNLHLMLNADKEFLAQKSAPHRDFYNVR 386 
QY	125 KVDTVHHSCMQKHILLRIKSKLRKEPDENVIFRDGTVMYLKEVFESLDLTGYDLND 184 
Db	387 KVDTVHHSCMQKHILLRIKSKLRKEPDENVIFRDGTLYTLREVFPESLDLTGYDLND 446 
QY	185 LLVDHADKSTFHFRDKFNKLKNPCGQSRLREI FLKQDNLIQGRFLAEITKQVFSDLASK 244 
Db	447 LLVDHADKSTFHFRDKFNKLKNPCGQSRLREI FLKQDNLIQGRFLGEITKQVFSDEASK 506 
QY	245 YQMAEYRISITYGRKQSEWDQLASWIVNNELHSGNVWMVLQIPRLYVVKEMGI VTSFONL 304 
Db	507 YQMAEYRISITYGRKQSEWDQLASWIVNNDLYSENVMVLIQPLRYNTIKDMGI VTSFQNI 566 
QY	305 LDNI FVLPFEVTIDPASHPOLHVFCLKOVVGGLVDDES KPERRPPTKGMPTEQMTNVNP 364 
Db	567 LDNI FLPFEATVDDSHQPQHVFCLKOVVGFDLVDES KPERRPPTKGMPPTAQMTNAFP 626 
QY	365 AFSYAYICYANLFTLNKLRESKGMTTIKFRPHAGEAGDVHDLAATFLLCNISHGINLR 424 
Db	627 AFSYVYYVCYANLVVNKLRESKGMTTITLRPHSGEAGDI DHAATAFLTCHSIAHGINLR 686 
QY	425 KSPVLQVLYLGOIGLAMSPLSNNSLFLDYHRNPFPPTFFOGLNVSLSTD DPLOIH LTKE 484 
Db	687 KSPVLQVLYYLAQIGLAMSPLSNNSLFLDYHRNPFPFPFLRGLNVSLSTD DPLOIH LTKE 746 
QY	485 PLVEEYSIAASLWKLSLCDICEIARNVS YQSGFSHALKAHWIGKNYFKRGPGAGNDIHRTN 544 
Db	747 PLVEEYSIAASVWLKSCADCLCEIARNVS YQSGFSHALKSHWIGKYKRGPDGDNDIHKTN 806 
QY	545 VPHIRVOFREMIWRNEKMLVYSDNEILLPDEL 576 
Db	807 VPIRVBFERDITWKEEQVYL-GKAVISDSV 837 

QNGXGO	PRELIMINARY;	PRT;	743 AA.
QNGXGO;			
AC			
DT	01-OCT-2000 (T-EMBLrel. 15, Created)		
DT	01-OCT-2000 (T-EMBLrel. 15, Last sequence update)		
DT	01-MAR-2003 (T-EMBLrel. 23, Last annotation update)		
DE	AMP deaminase.		
DE	AMPA.		
GN	Dictyostelium discoideum (Slime mold).		
OS	Eukaryota; Mycetozoa; Dictyostelid; Dictyostelium.		
OC	NCBI_TaxID=44689;		
OX	[1]		
RN	SEQUENCE FROM N.A.		
RP	Chae S., Iranfar N., Loomis W.F.;		
RA	"Dictyostelium discoideum AMP deaminase (amdA).";		
RL	Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.		
RT	EMBL; AF238311; AAF65407.1;		
DR	InterPro; IPR001365; A/AMP_deaminase.		
DR	InterPro; IPR006650; A/AMP_deam_sub.		
DR	InterPro; IPR006329; AMP_deaminase.		
DR	Pfam; PF00962; A_deaminase; 1.		
DR	TIGRFRAMS; TIGR01429; AMP deaminase; 1.		
DR	PROSITE; PS00485; A_DEAMINASE; 1.		
DR	SEQUENCE 743 AA; 85585 MW; 228C6E6B052BD73A CRC64;		
QY	Query Match	58.0%; Score 1798; DB 5; Length 743;	
Db	Best Local Similarity	59.1%; Pred. No. 1.2e-136;	
Db	Matches 337; Conservative		
QY	8 WEKEVINDPCTPKENPNPFYVPEPK--SEHVQTVDGVIVHYA--SEHKDCTESYYPVAD 62		
Db	50 WKADA---PVGEKPPSPFDESSTNCATEHMFKEVNGVYFVYNETDKMKNKALFVPH 106		
QY	63 A-TFFFTDLHVILKRVTAAGNTRTVCHNRKLNLLHKKFKFHLMLNADREFLAQKTAPHRDFY 121		
Db	107 TLAASYKDINNLMLSSYGAKFTFKRLQLLESKFNMHTLNDLSLELFQOKTAPHRDFY 166		
QY	122 NVRKVDTHVHHSACMNQKHLLRFKTSKLKKEPDEWIPRDGTYMTLKVFVPSLDLTGYDL 181		
Db	167 NVRKVDTHVHSSMNQKLLKFKRLKENPBEIVFRDDKYLTAEVFKYSLNLDVDEL 226		
QY	182 NVDLLDVHAKSTFHRFDKFNKYNPCGOSLRIFLKQDNLIQGRFLAELTKQVFSDL 241		
Db	227 SVDTLDVHADNNTFHRFDKFNKYNPCGOSLRIFLKTDLNLKGYLAEISKVFPTDLE 286		
QY	242 ASKYQMAEYRISIIYGRKQSEWDOLASMTVNNELSGNVVWLVIQIPRLYNVYKEMGIVTSF 301		
Db	287 SSKYQCAEYRLSIYGRKQSEWDTLASIVNDLSTFKVRWLIQVRLYDVVRETS-TYTF 345		
QY	302 QNLNDINFPVLPFEVITDPASHQPLHVLKQVVGDLVDDESKPERRPTKHMPTPQWTVN 361		
Db	346 QDFLANNVFPVLPFEVTKDPSSHPKHLFLQVVGIDCVDESKFEKFTKEKFPVPGWSSE 405		
QY	362 FNPAFSYAYCYANLFTLNLKRESKGMTTKFRPHAGEAGDVHDLAATFLCHNISHGI 421		
Db	406 HNPPYTYLYLYANLYTLNQFREEKGLNLTLPHPHGEAGEVDHMGAAFYLAHGINHI 465		
QY	422 NLKESPVLYLYLGOGLGAMSPLNNSLFDYHNRNPPPTFFORGLNVSLSTDDPLQLHL 481		
Db	466 NLRKTPVLYLYLTQIGAMSPUNNSLFTYINRNPFAFFARGLNVSISTDDPLQPHY 525		
QY	482 TKEPLVEEYSTAASLWLSSCDLCEIARNVSYQSGFHALKAHWIGKGYFKRGPAGNDI 541		
Db	526 TKEPLMEEYSTAQWRLSVCDICEIARNVLSQSGFEHNVKSHWLGPDYANS--GNDIK 583		
QY	542 RTNVPHIRVQFREMWRNEMKLVYSDNEIL 571		
Db	584 KTNISDIRVCFRNETLIEELHLILKSLQTL 613		
RESULT 3			
Q9DBT5	PRELIMINARY;	PRT;	798 AA.
ID	Q9DBT5		
AC	Q9DBT5		
AC	Q9DBT5;		

RESULT 2  
O9NGX0



```

DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 1200014F01RIK protein.
GN 1200014F01RIK
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=lung;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Mateda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Cariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
DR EMBL; AK004759; BAB23540.1; -
DR MGD; MGI:1921388; 1200014F01RIK.
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP deam_sub.
DR InterPro; IPR006329; AMP deaminase.
DR Pfam; PF00962; A deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
DR SEQUENCE 798 AA; 92023 MW; 2BC4F37E4006C7D5 CRC64;
SQ
Query Match 54.4%; Score 1687.5; DB 11; Length 798;
Best Local Similarity 54.2%; Pred. No. 1.2e-127;
Matches 318; Conservative 84; Mismatches 162; Indels 23; Gaps 6;
QY 1 PRVRVAPWEKEVINPCTPKPNPFTYVPEPKSEHVFTQVDGVIHYVA----DKDCTES 56
DB 220 PALEQHPYEH-----CEPSAMPGDLGL-----GLRMVGVVHVYTRRDPDCHCEV 265
QY 57 IYPVADATTFDTLHYLRVTAAGNTRTVCHNRLNLLHKKFKFHLMLNADREFLAQKTAP 116
DB 266 ELFPYDLQEFVADVNVLMALIINGPIKSCYRRLQYLSSKFQHVLLNEMKELAAQKVP 325
QY 117 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPDEWVIFRDGYMTLKVFESLDL 176
DB 326 HRDFYNIRKVDTHHASSCMNQKHLRFIKRAMKRHLREIIVHVEQREQLTRFEVESNNL 385
QY 177 TGYDLNVLLDVAHDKSTFHRFDKFNLYNCPGOSRLREIFLKQDNLIQGRFLAELTKQV 236
DB 386 TAYDLSVDTLDVHADNTRFHRFDKFNLYNCPGOSRLREIFLKQDNLIQGRFLAELTKQV 445
QY 237 FSDLASQYOMAEYRISYIGRKQSEWDQLASWIVNNELHSGNVVWLVOIPLRYNVMKMG 296
DB 446 MADLSESKYQNAELRLSYIGRDEWDKRLARWAVNKHSPNVRVLVQVPRFLFDVYRTKG 505
QY 297 IVTSFQNLNDNI FVPLFEVITDPASHPOLHVLKQVVGDLVDDESKPERRP-TKHMTPT 355
DB 506 QLANFQEMLENI FLPFEATVHPASHPELHLFLEHVDFGSDVDDSKPENHVFNLESPLP 565
QY 356 EQWTVNFPAPSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHUAATFLFLCH 415
DB 565 EAWTEEDNPPYAYLYYTFANMAMNLHRRQGRFHTVLRPHPCGEAGPIHHLVSFAFMLAE 624

```

```

DB 566 EAWTEEDNPPYAYLYYTFANMAMNLHRRQGRFHTVLRPHPCGEAGPIHHLVSFAFMLAE 625
QY 416 NISHGINLRKSPVLQYLYLQIGIAMSPLSNNSLFLDYHRNPFTTFFQGRGLNYSLSLSTDD 475
DB 626 NISHGLLRKAPVLQYLYLAQIGIAMSPLSNNSLFLSYHNPPLPEYLSRGLWSLSLSTDD 685
QY 476 PLQHLTKPELVEBSYIAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIKGNVFRKGP 535
DB 686 PLQHFHTKEPLMEYSIATQVWKLSSCDLCEIARNSVYQSGFSHALKAHWIKGNVFRKGP 745
QY 536 AGNDIHRTNVPHIRVQPREMIWRNEMKLVYS--DNEIL--IPDELDEL 578
DB 746 EGNDIRTNVEDIRVGRYETLCOELALITQAVQSEMLETIPEEVGI 792
RESULT 4
Q96IAL PRELIMINARY; PRT; 798 AA.
ID Q96IAL;
AC Q96IAL;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Adenosine monophosphate deaminase 2 (Isoform L).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC007711; AAH07711.1; -
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP deam_sub.
DR InterPro; IPR006329; AMP deaminase.
DR Pfam; PF00962; A deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
DR SEQUENCE 798 AA; 92070 MW; A31876E9F8C93EB CRC64;
SQ
Query Match 54.2%; Score 1680.5; DB 4; Length 798;
Best Local Similarity 54.3%; Pred. No. 4.3e-127;
Matches 317; Conservative 82; Mismatches 162; Indels 23; Gaps 6;
QY 1 PRVRVAPWEKEVINPCTPKPNPFTYVPEPKSEHVFTQVDGVIHYVA----DKDCTES 56
DB 219 PALEQHPYEH-----CEPSTMPGDLGL-----GLRMVGVVHVYTRRDPDCHCEV 264
QY 57 IYPVADATTFDTLHYLRVTAAGNTRTVCHNRLNLLHKKFKFHLMLNADREFLAQKTAP 116
DB 265 ELFPYDLQEFVADVNVLMALIINGPIKSCYRRLQYLSSKFQHVLLNEMKELAAQKVP 324
QY 117 HRDYNVRKVDTHVHSAACMNQKHLRFIKSKLRKEPDEWVIFRDGYMTLKVFESLDL 176
DB 325 HRDFYNIRKVDTHHASSCMNQKHLRFIKRAMKRHLREIIVHVEQREQLTRFEVESNNL 384
QY 177 TGYDLNVLLDVAHDKSTFHRFDKFNLYNCPGOSRLREIFLKQDNLIQGRFLAELTKQV 236
DB 385 TAYDLSVDTLDVHADNTRFHRFDKFNLYNCPGOSRLREIFLKQDNLIQGRFLAELTKQV 444
QY 237 FSDLASQYOMAEYRISYIGRKQSEWDQLASWIVNNELHSGNVVWLVOIPLRYNVMKMG 296
DB 445 MSDLSESKYQNAELRLSYIGRDEWDKRLARWAVNKHSPNVRVLVQVPRFLFDVYRTKG 504
QY 297 IVTSFQNLNDNI FVPLFEVITDPASHPOLHVLKQVVGDLVDDESKPERRP-TKHMTPT 355
DB 505 QLANFQEMLENI FLPFEATVHPASHPELHLFLEHVDFGSDVDDSKPENHVFNLESPLP 564
QY 356 EQWTVNFPAPSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHUAATFLFLCH 415
DB 565 EAWTEEDNPPYAYLYYTFANMAMNLHRRQGRFHTVLRPHPCGEAGPIHHLVSFAFMLAE 624

```

```
QY 416 NISHGINLRKSPVLYLYLGOIGLAMSPLSNLSFLDYHRNPFTFFQFQRLNVLSTDD 475
Db 625 NISHGLLRKAPVLYLYLGOIGLAMSPLSNLSFLSYHRNPFLPEYLSRGLMVSISTDD 684
QY 476 PLQIHLTKPLVEEYSIAASLWKLSSCDLCEIARNVSVOGFSHALKAHWIKNYFKRGPGAGNDIHRNVP 535
Db 685 PLQFHTKEPLMEYSIAATQVWKLSSCDLCEIARNVSVOGFSHALKAHWIKNYFKRGPGAGNDIHRNVP 744
QY 536 AGNDIHRNVPDIRVGRVYETLCQELALITQAVQSEMLETIPEE 575
Db 745 EGDNRRTNVPDIRVGRVYETLCQELALITQAVQSEMLETIPEE 788

RESULT 5
Q9P3N4 PRELIMINARY; PRT; 1008 AA.
ID Q9P3N4
AC Q9P3N4
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable AMP deaminase.
GN B7F18.150.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
ON NCBI_TaxID=5141;
RN [1]_TaxID=5141;
RP SEQUENCE FROM N.A.
RA Schulte U., Algen V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL389891; CAB97316.2; -
DR InterPro; IPR001365; A/AMP deaminase.
DR InterPro; IPR006650; A/AMP deam. sub.
DR InterPro; IPR006329; AMP deaminase.
DR Pfam; PF00862; A_deaminase; 1.
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.
DR PROSITE; PS00485; A_DEAMINASE; 1.
SQ SEQUENCE 1008 AA; 11317 MW; 6449671AFAE3F9B CRC64;

Query Match 52.9%; Score 1641; DB 3; Length 1008;
Best Local Similarity 54.8%; Pred. No. 9.2e-124;
Matches 309; Conservative 93; Mismatches 150; Indels 12; Gaps 2;

QY 10 KEVINDPCTPKPNPFTYVPEKSEHFQTVGVHVVADKDCSTSIYPVADATT---F 66
Db 332 EDFVLEDLLVPVGPDDFTYKLDE-----NGVVQIYSNDQHKADQPAIKIPTLREY 382
QY 67 FTDLHVLRTAAGNTRTVCHNRLNLEHKFKPHLMNADREFLAQKTAHPDRFYNVKRV 126
Db 383 YLDLEAILSVSDGPKSFAFRRLQVLEKFRDLYLLEGVDETADCKVPHRDFYNVKRV 442
QY 127 DTHVHSACNQHARLFTKSLRKEPDEWVIFRDCQTYMTLKFVFSLDLTGVLDNVDLL 186
Db 443 DTHVHSACNQHARLFTKSLRKEPDEWVIFRDCQTYMTLKFVFSLDLTGVLDNVDLL 502
QY 187 DVHADKSTHFRDFKFLKYNPCQSRRLRIFLQDNLIOGRFLAELTKQVFSLSASKYQ 246
Db 503 DMEHTDSFHRDFKFLKYNPCIGESRLRTIFLTKDNFNGRYLAETKEVIADLESKKYQ 562
QY 247 MARYISIVGRKQSEWDLQASWLVNNEHSGNVVWLQVIRLVNNTKGVGIVTSFQNLID 306
Db 563 MVEWRVSIYKSLDNDKLASWVDNKLFSHNVRVLVQVIRLVNNTKGVGIVTSFQNLID 622
QY 307 NIEVPLFEVITDPASHPOLHVLKQVGLDVLVDESKPERRPTKHMPTPEQNTVNPAP 366
Db 623 NIEPFLFEVITDPASHPOLHVLKQVGLDVLVDESKPERRPTKHMPTPEQNTVNPAP 682
QY 367 SYVAYCYANLFTLNKLRSGKMTTIKFRPHAGEAGDVHDLAATFLLCHNISHGINLRKS 426
```

```
Db 683 SYWYLYANWASLNHWKRRGFNTFVLRPHCGAGDSEHLAVALCCHSISHGLLRKV 742
QY 427 PVLYLYLGOIGLAMSPLSNLSFLDYHRNPFTFFQFQRLNVLSTDDPLQIHLTKPL 486
Db 743 PVLYLYLGOIGLAMSPLSNLSFLDYHRNPFTFFQFQRLNVLSTDDPLQIHLTKPL 802
QY 487 VVEYSIAASLWKLSSCDLCEIARNVSVOGFSHALKAHWIKNYFKRGPGAGNDIHRNVP 546
Db 803 IEEYAVAAQIYKLSFVDMCELAKNSVQSGYVERISKEQMLGKGNFDPGAQNTVTKTNP 862
QY 547 HIRVOFREMIRWNEKLVYSNEI 570
Db 863 DRBEFRYRTLEERGMYRYNAL 886

RESULT 6
Q9VY76 PRELIMINARY; PRT; 652 AA.
ID Q9VY76
AC Q9VY76
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG32626 protein.
GN CG32626 OR CG11058 OR CG11065 OR CG15762.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxID=7227;
RN [1]_TaxID=7227;
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang X., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Palazzolo M., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svendsen R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195 (2000).
RN [2]
```

RP SEQUENCE FROM N.A.  
RA Celniker S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,  
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,  
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
RA Pacleeb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,  
RA Williams S.M., Zaveri J.S., Smith H.O., Venter C.J., Rubin G.M.;  
RT "Sequencing of Drosophila melanogaster genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,  
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,  
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
RT "Annotation of Drosophila melanogaster genome."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA FlyBase;  
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AS003493; AAF09329.2;  
DR FlyBase; Fgn0052626; CG32626.  
DR InterPro; IPR001365; A/AMP\_deaminase.  
DR Pfam; PF00962; A\_deaminase; 1.  
DR PROSITE; PS00485; A\_DEAMINASE; 1.  
SQ SEQUENCE 652 AA; 76054 MW; 846D8012F2E277E2 CRC64;  
Query Match 52.8%; Score 1638; DB 5; Length 652;  
Best Local Similarity 53.9%; Pred. No. 8.e-124;  
Matches 305; Conservative 87; Mismatches 152; Indels 22; Gaps 5;  
7 PWEKEVINDPCTKPNPNFTVVPKSEHVQTVGVHVA-DKDCETESYPVADATT 65  
96 PWNVEF-----PNDEDFKTKP-----LNGVPHIYENDESSEIKYEYDMSQ 137  
66 FTFDLHYILRVTAAGNTRTVCHNRLNLLBHKFKHMLNADREFLAQKTAPHRDFVYVRK 125  
138 FVNDQVMCMNTADGLPKSCFYRRLCYLSSKYQMVLNRLNLAQKAVPHRDYNTRK 197  
126 VDVTHVHSAQMKQHLRLRFIKSLRKEPDEVVIFRDGYTMTLKEVFESLDLTGYDLNVDL 185  
198 VDVTHIAASQMKQHLRLRFIKSLRKEPDEVVIFRDGYTMTLKEVFESLDLTGYDLNVDL 257  
186 LDVHADKSTFHRFDKENLKNYFCGOSRLREIFLKDNLQGLFLAELTKQVPSDLSASKY 245  
258 LDVHADRNFTFHRFDKFNKYNPIGESRLREIFLKDNLQGLFLAELTKQVPSDLSASKY 317  
246 QMAEYRISYIGKQSEWDQLASVNNLHSGNVVWLVOIPLRYNNYKEMGVITSPONLL 305  
318 QNAELRLSYIGKSPDEYKLAQWADNDVSYNNIRLWLIQPLRFDIFKSNQMKQFQEL 377  
306 DNIPLVPLEFTVDPASHPOLHFLKQVGLDLDVDDSKPERRP--TKHMTPEQWNTVFN 363  
378 NNIFLPLPEATAPSKHPHRLFLQVIGFSDVDDSKPE-NPLFDNDVPRPEWNTYEN 436  
364 PAFSYAYCYANLFTNLKLRSGKMTTKIFRPHAGEAGDVHDLAATFLCLNISHGINL 423

Db 437 PPAYIYYVYANMTVLNKFQSRNMNTFVLRPHCGEAGPVQHLVCGFLMAENISHGLLL 496  
QY 424 RKSPVLQYLYLQIGTGLAMSPLSNNSFLDYHRNPFPTFFQFGLNVLSTDDPLQIHLTK 483  
Db 497 RKVPVLQYLYLTQIGIAMSPLSNNSFLNHRNPLPEYLARGLIISLSTDDPLQFHFHTK 556  
QY 484 EPLVEYSIAASLWKLSSCDLCEIARNVSYGSGSHALKAHWIKNYKPYKPGAGNDIIRT 543  
Db 557 EPLMEYSIAAQVWKLSSCDMCEIARNVSMGSPFHAIKQOQWLGPYYVEDIMGNDIIRT 616  
QY 544 NVPHRVQFREMWRNEMKLVYSNE 569  
Db 617 NVPEIRVAYRYETLDELUSNIFKVNQ 642  
RESULT 7  
Q96107 PRELIMINARY; PRT; 707 AA.  
ID Q96107;  
AC Q96107;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)  
DE GH10492p (CG32626-PC)  
GN CG32626 OR CG11058 OR CG11065 OR CG15762.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nunoo J., Pacleeb J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
EX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,  
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
RA Flogel C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.B., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleeb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., WoodageT, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.:  
 RT "The genome sequence of *Drosophila melanogaster*;"  
 RL Science 287:2185-2195 (2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,  
 RA Bantz J., An H., Baldwin D., Banzone J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferrera S., Frise E., Galle R.F., Gaig N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of *Drosophila melanogaster* genome;"  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Krommiller B., Marshall B., Milburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of *Drosophila melanogaster* genome;"  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY051429; AAK92853.1; -;  
 DR EMBL; AE003493; AAN09337.1; -;  
 DR FlyBase; FBgn0052626; CG32626.  
 DR InterPro; IPR001365; A/AMP\_deaminase.  
 DR InterPro; IPR006650; A/AMP\_deam sub.  
 DR InterPro; IPR006329; AMP\_deaminase.  
 DR Pfam; PF00962; A\_deaminase; 1.  
 DR TIGRFAMs; TIGR01429; AMP\_deaminase; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; 1.  
 SQ SEQUENCE 707 AA; 82035 MW; 0FE03C647022051A CRC64;  
 Query Match 52.8%; Score 1638; DB 5; Length 707;  
 Best Local Similarity 53.9%; Pred. No. 9.9e-124;  
 Matches 305; Conservative 87; Mismatches 152; Indels 22; Gaps 5;  
 QY 7 PWEKEVINDECTPKPNPNPTTYPEPKSEHVFTVDGVHIVYA-DKDCETESIPVADATT 65  
 DB 151 PWNVEP-----PNDEDPKPKP-----LNGVPHIYENDDESSEIKYEYDMSQ 192  
 QY 66 FFTDLHILRVTAAGNTRTRCHNRLNLEHKKFHLMLNADREFLAQKTAPHRDFYNVK 125  
 DB 193 FVNDQMVMCMNIADGPKSKFCYRRLCYLSKYQHVLLNRLRLAQKAVPHRDFYNTK 252  
 QY 126 VDTVHHSACMKNQHLRLFKSKLRKEPDEWIFRDGTVMTLKEVPESLDLTGYDLNVDL 185  
 DB 253 VDTTHRAASCNMKNQHLRLFKSKLRKEPDEWIFRDGTVMTLKEVPESLDLTGYDLNVDL 312  
 QY 186 LDVHADKSTHTRFDKFNKLNKPNCGSRLRIFLKDNLIOGRFLAELTKOVFSDLSASKY 245  
 DB 313 LDVHADRNTRHTRFDKFNKSNKPNIGESRLREVFLKTDNLNGKYFAQIIEKVAFDLESKY 372

QY 246 QMAEYRISYGRKQSEWDLASWIVNNELHSGNVVWLVIQIPRLNIVYKEMGIVTSFQNL 305  
 DB 373 QNAELRLSIYKSPDEWYKLAKWAINDVSSNIRLWLIQIPRLDFIKSNOMKSFQIL 432  
 QY 306 DNIIVPLFEVITIDPASHPOLHVLKQVGLDLDVDESKPERRP--TKHMPPTPEQWTVN 363  
 DB 433 NNIPLPLFEATARPSKHPHFLQYVIGFVDSVDESKPE-NPLFDNDVPRPEEWTYEEN 491  
 QY 364 PAFSYAYCYANLFTLNKLRKSGMTTIKFRPHAGEAGDVVDHLAATFLLCNTHSHGINL 423  
 DB 492 PPAYIYIYVYVNTVNLNFKQSRNMTFVLRPHCGEAGPVQHLVCGFLMAENISHGLLL 551  
 QY 424 RKSPLQYLYLQIGQIGLAMSPLSNLSFLDYHRNPFFTFQFQGLNVLSLSTDDPLQIH 483  
 DB 552 RKVPVLQYLYLTQIGIAMSPLSNLSFLNHRNPFLPEYLARGLIISLSLSTDDPLQHF 611  
 QY 484 EPLVEYSIAASLWKLSSCDLCETARNVSYOGSGSHALKAHWIKYKRGKPGAGNDIHT 543  
 DB 612 EPLMEYSIAAQVWKLSSCDLCETARNVSYOGSGSHALKAHWIKYKRGKPGAGNDIHT 543  
 QY 544 NVPHIRVOFREMIVRNEMKLVYSNDE 569  
 DB 672 NVPEIRVAYRIETLLDELSNIFKVNQ 697

## RESULT 8

Q18463  
 ID Q18463 PRELIMINARY; PRT; 865 AA.  
 AC Q18463;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)  
 DE Hypothetical 99.5 kDa protein.  
 GN C34F11.3  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2018 (1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bentley D.;  
 RT "The sequence of *C. elegans* cosmid C34F11.";  
 RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U46753; AAA85762.1; -;  
 DR WormPep; C34F11.3a; CE04164.  
 DR InterPro; IPR001365; A/AMP\_deaminase.  
 DR InterPro; IPR006650; A/AMP\_deam sub.  
 DR InterPro; IPR006329; AMP\_deaminase.  
 DR Pfam; PF00962; A\_deaminase; 1.  
 DR TIGRFAMs; TIGR01429; AMP\_deaminase; 1.  
 DR PROSITE; PS00485; A\_DEAMINASE; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 865 AA; 99499 MW; EA22B7D62D12EDFD CRC64;

Query Match 50.6%; Score 1569; DB 5; Length 865;  
 Best Local Similarity 53.5%; Pred. No. 5e-118;  
 Matches 303; Conservative 88; Mismatches 155; Indels 20; Gaps 7;



DR EMBL; BC040366; AAH40366.1; --  
SQ SEQUENCE 766 AA; 88652 MW; 94991133B18B0AA5 CRC64;

Query Match 49.6%; Score 1537; DB 11; Length 766;  
Best Local Similarity 51.8%; Pred. No. 1.6e-115;  
Matches 288; Conservative 91; Mismatches 163; Indels 14; Gaps 4;

QY 19 PKENPNPFYVPEPKS-EHVFQTVGVGVIHVYADKQCTE---SIVPVADATFTFFDLHYI 73  
DB 201 PUPQEDPYCLDDAPPNIGVLRMHGGLVFNQNTMLERQEPHSUPYPLETIYVDMSHI 260

QY 74 LRVTAAGNTRTYCHNRLNLEHKKFHLMLNADREFLAQTAHPHRDFYNNRVKVDTHVHRS 133  
DB 261 LALITDGPITYCHRELNFLESKFSLEHMLNEMSEFKLSNPHRDFYNNRVKVDTHIAA 320

QY 134 ACNQKHLLRFIKSKURKEPDEVVIFPDGTMTLKEVPESLDLTGVLDNLVDLVDHAKS 193  
DB 321 ACNQKHLLRFIKHTYTQTEPDRTVAEKLGKITLRQVDFSLHMDPYDLTVDSLDVHAGRQ 380

QY 194 THERFDKFNLYNPPCCQSRLREIFLKQDNLIOGRFLAELTKQVFSLSASKYQMAEYRIS 253  
DB 381 THERFDKFNLYNPPVGASELRLDYLKTNLYGGEYFARMVKEVARELEDSKYQSEPRUS 440

QY 254 IYGRKQSEVDQLASWIVNNELHSGNVVWLVQIPRLYNNYKEMGIVTSFQNLNDNIFVPLF 313  
DB 441 IYGRSPKEWSSSLARWFIQHKVYSPNNRWIIQVPIYDIFRSKLLPNFGKMLNIFLPLF 500

QY 314 EVTIDPASHQPLHVLKQVGLDLDVDESKPERPTKHM-----PTPEQWTVNVPAPSY 368  
DB 501 KATINPDQHRELHLFLKYVTGFSVDDESK---HSDHMFSDKSPSPDLMTSEQNPPYSY 556

QY 369 YAYCYANLFTLNKLRSGMTTIKFRPHAGAGVDVHLAATFLCHNLSHGINKLRKSPV 428  
DB 557 YLYTYANIMVNLNRRERGLSTFLRPHCGEAGSITHLVSAFLTADNLSHGLLKKSPV 616

QY 429 LOYLYLGOIGLAMSPLSNLSFLDYHRNPFPTFFQRLNLSLSTDDPLQIHLTKPLVE 488  
DB 617 LOYLYLAQIPAMSPLSNLSFLEYSKNPLREFLHKLHLSLSTDDPMQPHYTKALME 676

QY 489 EYSIASLWKLSCDICEIARNVYSGFSHALKAHWIGNKVPKRGAGNDIHRNTVPHI 548  
DB 677 EYIAAQAQWKLSTCDLCEIARNVLSGLSHQKQKFLGQNYKKEGPEGNDIRKTNVAQI 736

QY 549 RVQPREMIWNEKLV 564  
DB 737 RMAFRYETLCNELSFL 752

RESULT 11  
Q922J7 PRELIMINARY; PRT; 481 AA.

ID Q922J7 AC Q922J7; DT 01-DEC-2001 (Tremblrel. 19, Created) DT 01-DEC-2001 (Tremblrel. 19, Last sequence update) DE 01-MAR-2003 (Tremblrel. 23, Last annotation update) DE Unknown (Protein for IMAGE:3591312) (Fragment). GN AMPD3.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC007183; AAH07183.1; --  
DR MGD; MGI:1096344; Ampd3.  
DR InterPro; IPR001365; A/AMP deaminase.  
DR InterPro; IPR006650; A/AMP deam\_sub.  
DR InterPro; IPR006329; AMP deaminase.  
DR Pfam; PF00962; A\_deaminase; 1.  
DR TIGRFAMs; TIGR01429; AMP deaminase; 1.  
DR PROSITE; PS00485; A\_DEAMINASE; 1.

FT NON TER 1  
SQ SEQUENCE 481 AA; 56376 MW; B6DF4F177F105A33 CRC64;

Query Match 46.0%; Score 1424.5; DB 11; Length 481;  
Best Local Similarity 55.5%; Pred. No. 1.1e-106;  
Matches 261; Conservative 77; Mismatches 123; Indels 9; Gaps 2;

QY 100 HMLNADREFLAQTAHPHRDFYNNRVKVDTHVHHSACNMQKHLLRFIKSKLRKEPDEVVIF 159  
DB 2 HEMLNEMSEFKELSGNPHRDFYNNRVKVDTHIAAACNMQKHLLRFIKHTYTQTEPDRTVAE 61

QY 160 RDGTMTLKEVPESLDLTGVLDNLVDLVDHAKDSTFHRFDKFNLYNPPCCQSRLREIFLK 219  
DB 62 KLGRKITLRQVDFSLHMDPYDLTVDSLDVHAGRQTFHRFDKFNLYNPPVGASELRLDYLK 121

QY 220 QDNLIQGRFLAELTKQVFSLSASKYQMAEYRISYIGRKQSEVDQLASWIVNNELHSGNV 279  
DB 122 TENYLGGEYFARMVKEVARELEDSKYQSEPRLSYIGRSPEKWSLSARWFIQHKVYSPNN 181

QY 280 VMLVQIPRLYNNYKEMGIVTSFQNLNDNIFVPLPEVTIDPASHQPLHVLKQVGLDLDV 339  
DB 182 RWIQVPIYDIFRSKLLPNFGKMLNIFLPLFKATINPDQHRELHLFLKYVTGFSVD 241

QY 340 DESKPERPTKHM-----PTPEQWTVNVPAPSYAYCYANLFTLNKLRSGMTTIK 394  
DB 242 DESK---HSDHMFSDKSPSPDLMTSEQNPPYSYLYTYANIMVNLNRRERGLSTFL 297

QY 395 RPHAGEAGDVHLLAATFLCHNLSHGINKRSPVLYLYLGOIGLAMSPLSNLSFLDY 454  
DB 298 RPHCGEAGSITHLVSAFLTADNLSHGLLKKSPVLYLYLAQIPAMSPLSNLSFLLEY 357

QY 455 HRNPPTFFQRLNLSLSTDDPLQIHLTKPLVEYSIAASLWKLSCDICEIARNVYVQ 514  
DB 358 SKNPLREFLHKLHLSLSTDDPMQPHYTKALMEYIAAQAQWKLSTCDLCEIARNVQL 417

QY 515 SGFSHALKAHWIGNKVPKRGAGNDIHRNTVPHI RVQPREMIWNEKLV 564  
DB 418 SGLSHQKQKFLGQNYKKEGPEGNDIRKTNVAQIRMAFRYETLCNELSFL 467

RESULT 12  
Q9IE53 PRELIMINARY; PRT; 672 AA.

ID Q9IE53 AC Q9IE53; DT 01-MAR-2003 (Tremblrel. 23, Created) DT 01-MAR-2003 (Tremblrel. 23, Last sequence update) DE 01-MAR-2003 (Tremblrel. 23, Last annotation update) DE AMP deaminase, putative (EC 3.5.4.6). GN MAL13P1.146.  
OS Plasmodium falciparum (isolate 3D7).  
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
OC NCBI\_TaxID=36329;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,  
RA Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,  
RA Ormond D., Sanders M., Hayes B., Hall S., Quail M., Barrell B.,  
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL844509; CAD52413.1; --  
KW Hydrolase.  
SQ SEQUENCE 672 AA; 79308 MW; F50A6CF4CEA62CCB CRC64;

Query Match 42.7%; Score 1325; DB 5; Length 672;  
Best Local Similarity 47.3%; Pred. No. 1.9e-98;  
Matches 268; Conservative 89; Mismatches 186; Indels 24; Gaps 8;

QY 8 WEKEVINDPCTPKPNPNPTVYVPEPKSEHVFQTVGVHVVAD-----KD-CTES- 56  
DB 109 YKSKNIDYSESSEPIYNPNVKNLKNCFNAFINFVDGIFVHWDPTDEGSSRDMCVESN 168

QY 57 -----IYPVADATFTFDLHYILRVTAAGNTRTYCHNRLNLEHKKFHLMLNADREFLAQ 112  
DB 169 KLANHRNTKSAEDYLSIQEIMNVVQDPAKSCFCQRLKYLEKFDHFMFNGPULSET 228



Best Local Similarity 37.5%; Pred. No. 8.3e-68;		Matches 169; Conservative 34; Mismatches 72; Indels 5; Gaps 3;	
Matches 206; Conservative 100; Mismatches 219; Indels 25; Gaps 8;			
QY	27 TVPEPKSHVFTQVDGVI---HVVADKDCSTSIYPVADATFTFDLHYI---LRVTAAG 80	QY	304 LLDNIFVPLREVTIDPASHPOLHVFLKQVGLDLDVDDSKERRP-TKHMPTPEOWTNVF 362
Db	1249 SYLQFDEDEWKYKTVGGIIVPHEVHQIPRLQPDQMYHTEFCNHBVEIRCLDNIIRV--- 1304	Db	1 MLENIFLFLFATVHPASHPELHLFLEHVDGFDSDVDDSKFENHVFNLESPLPEAWVEED 60
QY	81 NRTVTCHNLNLEHKKFPHLMNADREF--LAQKTAPHRDFVNVKVDTHVHSAQNMQ 138	QY	363 NPAFSYAYCYANLFTLNKLRRESKGMTTIKFRPHAGEAGDVHDHAAATFLLCNISHGIN 422
Db	1305 --RDFALRLQLLEHFRKLUHAAVNHSHRELGSAAASHNRDFYQSTKVDNNIRMETGMTA 1362	Db	61 NPPVAYLYITFANMAMLNHLRRQGFHTFVLRPHCGEAGPIHHLVSAFMAENISHGLL 120
QY	139 KHLLEIKSLRKEPDEVVIFRDGTY-MTLKEVFESLDLTGYDLNVLDLVDHADKSTFHR 197	QY	423 LRKSPVLYLYLGOIGLAMSPLNSLFLDYHRNPFTFFQRLGNVLSLSTDDPLQIHLT 482
Db	1363 RQLLAIFIVDKATHNGDDIVSHPKGEPQTLRQLLADLHITADSLTVDDLNVQAGATS --- 1419	Db	121 LRKAPVLYLYLAQIGIAMSPLNSLFLSYHRNPFLPEYLSRGLMVLSLSTDDPLQIHLT 180
QY	198 FKFNLKYNPCG-----QSRLEIFLQDNLIQGRFLAELTKQVFSDDLASKYQWAEYRI 252	QY	483 KEPLVEEYSIAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIGKNYFKRGPAGNDIHR 542
Db	1420 -SNGGAPQNPFASEGQOQDELLTLTKDNQMGRYFABELTKRTFEELSRDQHTFSESRL 1478	Db	181 KEPLMEEYSIATQVMKSSCDMLARNSVLMSGFSHKVKSHWLGPNYTKGPEGNDIHR 240
QY	253 SIYGRKQSEMDOLASWVNNELHSGNVVLVQIPRLYVYKEMGIIVTSFQNLNLDNIFVPL 312	QY	543 TNVPHIRVQFREMIRWNRNEMKLVYS--DNEIL--IPDELDL 578
Db	1479 SVTGASAEWALLSHWFDTHGSSSHNQWVQPRIYSIRKAGRVASFAEYLEHVFEP 1538	Db	241 TNVPDIRVGYRYETLCQELALITQAVQSEMLETIPPEVGI 280
QY	313 FEVTIDPASHPOLHVFLKQVGLDLDVDDSKERRPTEKHMPTPEOWTNVFNPAFSYAYY 372		
Db	1539 WRISLHPNSDPRLFHINIAAFDCVEDERRPDVPLHLMRSPHEWTTEDPEPPNYLYH 1598		
QY	373 CVANLFTLNKLRSGMTTIKFRPHAGEAGDVHDHAAATFLLCNISHGINLRKSPVLYL 432		
Db	1599 LYANLRLNCFQRRRFVSFRPSCGAGGVHDHLLIGGLAQSVNYGVRLADSAFLOYL 1658		
QY	433 YLGOIGLAMSPLNSN-LFLDYHRNPFTFFQRLGNVLSLSTDDPLQIHLTKPELVEEYS 491		
Db	1659 FYLAQIGVTLSPUSNNKQLNLYHRNPFPQFFRRGLRVSLGTDPSLLYHHTQEPLEEYS 1718		
QY	492 IAASLWKLSSCDLCEIARNSVYQSGFSHALKAHWIGKNYFKRGPAGNDIHRNTNPHIRVQ 551		
Db	1719 IASKIWKLSPNDLSEVARNVLLSNFSLRFEKELGAMHFLSSSAGNDVAKTHLSDRVVA 1778		
QY	552 FREMIWRNEM 561		
Db	1779 YRFEAYHTEV 1788		

Search completed: December 15, 2003, 13:28:07  
Job time : 43 secs

RESULT 15	
Q91YI2	PRELIMINARY; PRT; 286 AA.
AC	Q91YI2;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hypothetical protein.
GN	1200014F01RIK.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=FVB/N;
RA	Strausberg R.;
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC016662; AAH16662.2; -
DR	MGD; MGI:1921389; 1200014F01RIK.
DR	InterPro; IPR001365; A/AMP deaminase.
DR	InterPro; IPR006650; A/AMP deam_sub.
DR	Pfam; PF00962; A deaminase; 1.
DR	PROSITE; PS00485; A DEAMINASE; 1.
KW	Hypothetical protein.
SQ	SEQUENCE 286 AA; 32677 MW; 43106743E178C39D CRC64;
Query Match 29.1%; Score 901.5; DB 11; Length 286;	
Best Local Similarity 60.4%; Pred. No. 1e-64;	



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 15, 2003, 13:25:03 ; Search time 20 Seconds

(without alignments)  
2779.275 Million cell updates/sec

Title: US-10-019-633-2

Perfect score: 3100

Sequence: 1 PRVRVAPWEKEVINDPCTPK.....NEMKLVSNDNEILIPDEL 578

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: Pirl:\*

2: Pirl:\*

3: Pirl:\*

4: Pirl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2624.5	84.7	600	T01259	AMP deaminase homo
2	1680.5	54.2	760	A44313	AMP deaminase (EC
3	1647	53.1	810	S49744	AMP deaminase (EC
4	1641	52.9	924	T50996	probable AMP deami
5	1617	52.2	846	T32621	amp deaminase - fi
6	1569	50.6	865	T15771	hypothetical prote
7	1549	50.0	774	S68147	AMP deaminase (EC
8	1549	50.0	776	S68146	AMP deaminase (EC
9	1505	48.5	747	A27366	AMP deaminase (EC
10	1497	48.3	747	I39444	AMP deaminase (EC
11	1186.5	38.3	605	S59996	AMP deaminase (EC
12	627	20.2	888	S50801	AMP deaminase homo
13	625	20.2	797	S44546	probable membrane
14	350	11.3	88	A37056	AMP deaminase (EC
15	149	4.8	334	G97269	adenosine deaminas
16	138.5	4.5	387	T35436	probable adenosine
17	121.5	3.9	339	T11785	adenosine deaminas
18	120.5	3.9	352	G86660	adenosine deaminas
19	120	3.9	471	T41581	hypothetical coile
20	117.5	3.8	333	A64919	adenosine deaminas
21	117.5	3.8	333	C90920	adenosine deaminas
22	117.5	3.8	333	H85768	adenosine deaminas
23	115.5	3.7	3071	T20345	adenosine deaminas
24	114.5	3.7	435	H90578	vacuolar protein s
25	112.5	3.6	868	G71691	conserved hypothet
26	112.5	3.6	873	B86471	hypothetical prote
27	112	3.6	888	E82885	hypothetical prote
28	112	3.6	1984	A44396	P-type cation tran
29	111.5	3.6	316	H83625	probable adenosine

RESULT 1

T01259

AMP deaminase homolog F16M14.21 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 31-Mar-2003

C:Accession: T01259; A84803

R:Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes,

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.

A:Reference number: Z14213

A:Accession: T01259

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-600 <ROU>

A:Cross-references: ENBL:AC003028; NID:G3335356; PID:G3335375

A:Experimental source: cultivar Columbia

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84803

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-600 <STO>

A:Cross-references: GB:AE002093; NID:G3335375; PIDN:AAC27176.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g38280; F16M14.21

A:Map position: 2

A:Introns: 64/3; 82/1; 124/3; 184/3; 207/2; 255/1; 275/3; 313/3; 360/3; 411/3; 429/3; 46

C:Superfamily: AMP deaminase

Query Match 84.7%; Score 2624.5; DB 2; Length 600;

Best Local Similarity 84.5%; Pred. No. 2.9e-187;

Matches 475; Conservative 45; Mismatches 41; Indels 1; Gaps 1;

QY 5 VAPWEKEVINDPCTPKPNPFTYVPKSEHVFTVDGVHVVADKCTESIVPADAT 64

DB 34 VAPWEKEVINDPCTPKPNPFTYVPKSEHVFTVDGVHVVADKCTESIVPADAT 93

QY 65 TFFTDHLVILRTVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAAPHDRDFYVNR 124

DB 94 AFTFDLHVILRTVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAAPHDRDFYVNR 153

QY 125 KVDTHVHHSACMNQKHLRFIKSKLRKEPDSEVIFRDGTMTLKEVESFESLDTGYDLNVD 184

DB 154 KVDTHVHHSACMNQKHLRFIKSKLRKEPDSEVIFRDGTMTLKEVESFESLDTGYDLNVD 213

QY 185 LLDVHADKSTPHRFDKFNKLKYNPCQSRRLRIFLQKQNLIOGRFLAELTKQVFSDSLASK 244

DB 214 LLDVHADKSTPHRFDKFNKLKYNPCQSRRLRIFLQKQNLIOGRFLAELTKQVFSDSLASK 273



QY 96 KFKFHLMLNADREFLAQKTAHPRDFYNNVRKVDTHVHSSACMNQKHLRLRIKSKLRKEPDE 155  
Db 329 RWNLYLLNEYQETSVSXKNPHRDFYNNVRKVDTHVHSSACMNQKHLRLRIKSKLRKXDE 388  
QY 156 VIFRDTGTWTLKEVPESLDLTGYDLNVDLVDHDKSTFHRFDKFNKYNPCGOSRLRE 215  
Db 389 KVIKFDGKLLTLDEVSRLHLAGYDLSIDTLDMAHAKOTFHRFDKFNKYNPCGOSRLRE 448  
QY 216 IFKQDNLIQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDQASWIVNNELH 275  
Db 449 IFKTNMYIKGTYLADITKQVIFDLENSKYQCEYRISYGRSLDEWDKSLASWIDNKVI 508  
QY 276 SGNVWLVOIPLRYNNYKEMGIVTSFQNLDDNIFPLFEVTTDPAHPQHLVFLQVGL 335  
Db 509 SHNRWLVOIPLRYDIYKGTIGVQSDICKNLQPLFEVTKNPOSHPKHLVFLQVIGF 568  
QY 336 DLVDDSKDEPRPTKHPMTPEQWNTNVPNFAFSYAYCYANFTLNKLRKESGMMTIKFR 395  
Db 569 DSVDDSKVDPRFHRKYPKESLWEAPONPPYSLYLYSNVASLNQWRAKGFNTVLUR 628  
QY 396 PHAGAGVDHAAFLPFLCHNTSHGINKRSPVQYLYLGOIGLAMSPLSNNSLFLDYH 455  
Db 629 PHCGBAGDPEHLVSAVLLAHGISHGILLKRPVQVLYYLDQVGIAMSPLSNALFLTYD 688  
QY 456 RNPFTFFORGLNVLSTDDPQIHLTKPEPLVEEYSIAASLWKLSSCDLCEIARNVSQOS 515  
Db 689 KNPPFRYFKRGLNVLSTDDPQIHLTKPEPLVEEYSIAASLWKLSSCDLCEIARNVSQOS 515  
QY 516 GFSAHKAHWIGKNYFKRGPAGNDIHRNTVPHIRVOFREMWRNEMKLV 564  
Db 749 GHEAQIKHGWIGKDFDKSVGEVDVVRNTVPDIRINRYDTLSTLELV 797

## RESULT 4

T50996  
probable AMP deaminase [imported] - Neurospora crassa  
N/Alternate names: protein B7F18.150  
C/Species: Neurospora crassa  
C/Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 31-Mar-2003  
C/Accession: T50996  
R/Schulte, U.; Aign, V.; Hohelsel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, July 2000  
A/Reference number: Z25286  
A/Accession: T50996  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-924 <SCH>  
A/Cross-references: EMBL:AL398891; GSPDB:GN00116; NCSP:B7F18.150  
A/Experimental source: BAC clone B7F18; strain OR74A  
C/Genetics:  
A/Gene: NCSP:B7F18.150  
A/Map position: 6  
A/Introns: 795/2  
C/Superfamily: AMP deaminase

Query Match 52.9%; Score 1641; DB 2; Length 924;  
Best Local Similarity 54.8%; Pred. No. 7.6e-114;  
Matches 309; Conservative 93; Mismatches 150; Indels 12; Gaps 2;  
QY 10 KEVINDPCTPKNPFPFFIVPEPKSEHVQTVGVHIVYADKDCSTESYPVADATT---F 66  
Db 248 EFDVLLELLPVGPDPDFTYKLE-----NGVYQIYSDNQHKEADQPAIKIPTLEY 298  
QY 67 FDTLHVILRVTAAGNTRTVCHNRLNLEHKFKFHLMLNADREFLAQKTAHPRDFYNNVRKV 126  
Db 299 YLDLEAILSVSDGSKSAFRLQYLEKRFDLVLLGEGYDQACKKVPKPHRDFYNNVRKV 358  
QY 127 DTHVHSSACMNQKHLRLRIKSKLRKEPDEWIFRDTGTWTLKEVPESLDLTGYDLNVDLL 186  
Db 359 DTHVHSSACMNQKHLRLRIKSKLRKEPDEWIFRDTGTWTLKEVPESLDLTGYDLNVDLL 418  
QY 187 DVHADKSTFHRFDKFNKYNPCGOSRLREIFLKQDNLIQGRFLAELTKQVSDLSASKYQ 246

Db 419 DMHAHTDSFHRFDKFNKYNPCGOSRLREIFLKQDNLIQGRFLAELTKQVSDLSASKYQ 478  
QY 247 MAEYRISYGRKQSEWDQASWIVNNELHSGNVWLVOIPLRYNNYKEMGIVTSFQNLDD 306  
Db 479 MVEMRVSIIYKSLDEWDKSLASWVVDNKLFSHNRWLVOIPLRYNNYKEMGIVTSFQNLDD 538  
QY 307 NIFVPLFEVTTDPAHPQHLVFLQVGLDLDVDDSKDEPRPTKHPMTPEQWNTNVPNFAF 366  
Db 539 NIFVPLFEVTTDPAHPQHLVFLQVGLDLDVDDSKDEPRPTKHPMTPEQWNTNVPNFAF 598  
QY 367 SVYAYCYANFTLNKLRKESGMMTIKFRPHAGAGVDHAAFLPFLCHNTSHGINKRKS 426  
Db 599 SYWIIYLANWASLNHWRKRGFTFVLRPHCGAGDSEHLAAVLAALCCHSHGILLRKY 658  
QY 427 PVLOLYLYLGOIGLAMSPLSNNSLFLDYHRRNPFTFFORGLNVLSTDDPQIHLTKPEPL 486  
Db 659 PVLOLYLYLGOIGLAMSPLSNNSLFLDYHRRNPFTFFORGLNVLSTDDPQIHLTKPEPL 718  
QY 487 VEEYSIAASLWKLSSCDLCEIARNVSQOSGFSAHKAHWIGKNYFKRGPAGNDIHRNTVP 546  
Db 719 IEEYAVAAQIYKLSVDMCELAKNVSQOSGYERSIKEQWLGNFDPGAGNTVMVKTNPV 778  
QY 547 HIRVOFREMWRNEMKLVYSDNEI 570  
Db 779 DRBEFRTYTLLEBERGMVRYNAL 802

## RESULT 5

T39261  
amp deaminase - fission yeast (Schizosaccharomyces pombe)  
C/Species: Schizosaccharomyces pombe  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C/Accession: T39261  
R/Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Davis, P.; Churcher, C.M.  
submitted to the EMBL Data Library, August 1999  
A/Reference number: Z21840  
A/Accession: T39261  
A/Status: preliminary; translated from GB/EMBL/DBBJ  
A/Molecule type: DNA  
A/Residues: 1-846 <LYN>  
A/Cross-references: EMBL:AL110295; PIDN:CAB53720.1; GSPDB:GN00067; SPDB:SPBC106.04  
A/Experimental source: strain 972h-; cosmid c106  
C/Genetics:  
A/Gene: SPDB:SPBC106.04  
A/Map position: 2  
A/Introns: 13/2; 72/3  
C/Superfamily: AMP deaminase

Query Match 52.2%; Score 1617; DB 2; Length 846;  
Best Local Similarity 57.8%; Pred. No. 4.1e-112;  
Matches 304; Conservative 84; Mismatches 126; Indels 12; Gaps 3;  
QY 43 GVHIVYADKDC-----TESIYPVADATTFTDHLVILRVTAAGNTRTVCHNRLNLEHKFK 98  
Db 245 GIFQVYENDSAYIAGTSPFH-IPTIRYIIDLEFLLSASSDGPSSKFSFRRLQYLEGRWN 303  
QY 99 FHLMLNADREFLAQKTAHPRDFYNNVRKVDTHVHSSACMNQKHLRLRIKSKLRKEPDEWVI 158  
Db 304 MYMLNERYQELADTKVPHRDFYNNVRKVDTHVHSSALANQKHLRLRIKSKLRKEPDEWVI 363  
QY 159 FRDGTWTLKEVPESLDLTGYDLNVDLVDHDKSTFHRFDKFNKYNPCGOSRLREIFL 218  
Db 364 WRDCKFTLQEVFDSLAKTSYDLSIDTLDMAHAKOTFHRFDKFNKYNPCGOSRLREIFL 423  
QY 219 KQDNLIQGRFLAELTKQVSDLSASKYQMAEYRISYGRKQSEWDQASWIVNNELHSGN 278  
Db 424 KTDNDINGRYLAELTKQVSDLSASKYQMAEYRISYGRNREEMDKLAAWIDNLELSPN 483  
QY 279 VVWLVOIPLRYNNYKEMGIVTSFQNLDDNIFVPLFEVTTDPAHPQHLVFLQVGLD 338  
Db 484 VRWLIQVRLYDVYKKSIGI VETFEVVRNVPFLEVTKDPRTPHKLVHFLQVRIGFDSV 543



Db 547 KSPNDPWTSEONPPYSYLYMYANIMVNLRRERGLSTFLFRPHCGEAGSITHLVA 606  
QY 411 FLCHNISHGINLRKSPVLQYLYLQIGLAMSPLSNLSFLDYHRNPPTFFQRLNVS 470  
Db 607 FLTADNISHGLLKKSPVLQYLYLAQIPAMSPLSNLSFLSKNPLREFLHKLHVS 666  
QY 471 LSTDPLQHLKPELVEYSYTAASLWKLSSCDLCEIARNVYQSGFSAKAKHIGKNY 530  
Db 667 LSTDPMQPHYTKALMEBYATAAQWKLSTCDLCEIARNVLSQSLSHQKQKFLGQNY 726  
QY 531 FKRGAGNDIHRNTVPHIRVQFREMWRNEMKLV 564  
Db 727 YKEGPEGNDIRKTNVAQIRMAFRYETLCNELSFL 760

RESULT 8  
S68146  
AMP deaminase (EC 3.5.4.6), erythrocyte, splice form 1a - human  
N:Alternate names: AMP deaminase isoform E  
N:Contains: AMP deaminase splice form 1b  
C:Species: Homo sapiens (man)  
C:Date: 06-Dec-1996 #sequence revision 13-Mar-1997 #text change 20-Jun-2000  
C:Accession: S68146; S68148; A45071; B45071; S28149; S27955  
R:Mahmke-Zizelman, D.K.; Eddy, R.; Shows, T.B.; Sabina, R.L.  
Biochim. Biophys. Acta 1306, 75-92, 1996  
A:Title: Characterization of the human AMPD3 gene reveals that 5' exon usage is subject  
A:Reference number: S68146; MUID:96201708; PMID:8611627  
A:Accession: S68146  
A:Molecule type: DNA  
A:Residues: 1-776 <MAH>  
A:Cross-references: EMBL:U29925  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
A:Note: only a small part of the nucleic acid sequence is shown  
A:Note: only a small part of the translation is shown  
A:Note: splice form 1a  
A:Accession: S68148  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 10-776 <MAW>  
A:Cross-references: EMBL:U29925  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995  
A:Note: splice form 1b; Met-10 is the initiator  
R:Mahmke-Zizelman, D.K.; Sabina, R.L.  
J. Biol. Chem. 267, 20866-20877, 1992  
A:Title: Cloning of human AMP deaminase isoform E cDNAs. Evidence for a third AMPD gene  
A:Reference number: A45071; MUID:93015995; PMID:1400401  
A:Accession: A45071  
A:Molecule type: mRNA  
A:Residues: 1-216 <MA2>  
A:Cross-references: EMBL:M84720; NID:g178548; PIDN:AAAS8365.1; PID:g178549  
A:Note: splice form 1a (fragment)  
A:Note: sequence extracted from NCBI backbone (NCBIP:116076)  
A:Accession: B45071  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 10-776 <MA3>  
A:Cross-references: GB:M84721; NID:g178550; PIDN:AAA58366.1; PID:g178551  
A:Note: sequence extracted from NCBI backbone (NCBIP:116085)  
A:Note: splice form 1b; Met-10 is the initiator  
R:Yamada, Y.; Goto, H.; Ogasawara, N.  
Biochim. Biophys. Acta 1171, 123-128, 1992  
A:Title: Cloning and nucleotide sequence of the cDNA encoding human erythrocyte-specific  
A:Reference number: S28149; MUID:93042002; PMID:1420359  
A:Accession: S28149  
A:Molecule type: mRNA  
A:Residues: 10-776 <YAM>  
A:Cross-references: GB:D12775; NID:g219456; PIDN:BA02240.1; PID:g219457  
A:Note: splice-form 1b; Met-10 is the initiator  
C:Genetics:  
A:Gene: GDB:AMPD3  
A:Cross-references: GDB:136013; OMIM:102772  
A:Map position: 11p15-11p15  
A:introns: 8/1; 83/2; 151/3; 206/1; 279/2; 322/3; 387/3; 431/3; 486/2; 528/3; 583/2; 623

C:Superfamily: AMP deaminase  
C:Keywords: alternative initiators; alternative splicing; erythrocyte; hydrolase  
F:1-776/Product: AMP deaminase splice form 1a #status predicted <LSPL>  
F:10-776/Product: AMP deaminase splice form 1b #status predicted <SSPL>

Query Match 50.0%; Score 1549; DB 2; Length 776;  
Best Local Similarity 50.9%; Pred. No. 4.1e-107; Mismatches 175; Indels 14; Gaps 4;  
Matches 292; Conservative 93;

QY 1 PRVRVAPWEKVINDPCTPKENPNPFTYVPBPKS-EHVFQTVGVGVHVVYADKCTE---- 55  
Db 193 PRADTAPPEGLPDHPHPPPLQEDPYCLDDAPPNLDYLVMHMQGGLFYVKNKMLEHOEP 252  
QY 56 SIYPVADATFTFDLHYILRVTAAGNTRVTCNRLNLLHKKPKFHLMLNADREFLAQKTA 115  
Db 253 HSLPYPDLETYVDMSHILALITDGPRTKYCHRRINFLSKFSLHEMLNEMSEFKELKSN 312  
QY 116 PHRDFYNNRVKVDTHVHSHACMNQKHLRFKSKLRKEDEVVIFRDGYMTLKEVFESLD 175  
Db 313 PHRDFYNNRVKVDTHHAAACMNQKHLRFKHTYQTEPDRTVAEKRGRKITLRQVFDGLH 372  
QY 176 LTGYDLNVDLLDVHADKSTFHRFDKFNLYNPGQSRUREIFLQDNLIQGRFLAELPKQ 235  
Db 373 MDPYDLTVDSLVDVAGRTFHRFDKFNKYNPVGASELRDLYLKTENYLGGEYFARMVKE 432  
QY 236 VFSDLSASKYQWAEYRISYGRKSEWMDQLASWVNNELHSGNVVWLQVLPRLVNYKEM 295  
Db 433 VARELESKYQVSEPRLSYGRSPPEWPNLAYWFIQHKVYSPNMRWIIQVPRIVDFRSK 492  
QY 296 CIVTSFQNLNINIFVPLFEVITIDPASHPOLHVFILQVVGGLDVDDDESKPERPTKHM--- 352  
Db 493 KLLNFGKMLENIFLPLFKATINPDQHRELHLFLKYVTGFDSDVDESK---HSDHMFSD 548  
QY 353 --PPEQMTNVNPAFYSYVYCVANLFTLNKLRESKGMTTIKRPFHAGEAGDVDDHLAAT 410  
Db 549 KSPNPDVWTSQNPSPYSYLYMYANIMVNLRRERGLSTFLFRPHCGEAGSITHLVA 608  
QY 411 FLCHNISHGINLRKSPVLQYLYLQIGLAMSPLSNLSFLDYHRNPPTFFQRLNVS 470  
Db 609 FLTADNISHGLLKKSPVLQYLYLAQIPAMSPLSNLSFLSKNPLREFLHKLHVS 668  
QY 471 LSTDPLQHLKPELVEYSYTAASLWKLSSCDLCEIARNVYQSGFSAKAKHIGKNY 530  
Db 669 LSTDPMQPHYTKALMEBYATAAQWKLSTCDLCEIARNVLSQSLSHQKQKFLGQNY 728  
QY 531 FKRGAGNDIHRNTVPHIRVQFREMWRNEMKLV 564  
Db 727 YKEGPEGNDIRKTNVAQIRMAFRYETLCNELSFL 762

RESULT 9  
A27366  
AMP deaminase (EC 3.5.4.6), skeletal muscle - rat  
N:Alternate names: adenylic acid deaminase; AMP aminase; myoadenylate deaminase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999  
C:Accession: A27366; I57509  
R:Sabina, R.L.; Marquetant, R.; Desai, N.M.; Kaletka, K.; Holmes, E.W.  
J. Biol. Chem. 262, 12397-12400, 1987  
A:Title: Cloning and sequence of rat myoadenylate deaminase cDNA. Evidence for tissue-sp  
A:Reference number: A27366; MUID:87308255; PMID:3624265  
A:Accession: A27366  
A:Molecule type: mRNA  
A:Residues: 1-747 <SAB>  
A:Cross-references: GB:J02811; NID:g202882; PIDN:AA854086.1; PID:g202883  
R:Mineo, I.; Clarke, P.R.H.; Sabina, R.L.; Holmes, E.W.  
Mol. Cell. Biol. 10, 5271-5278, 1990  
A:Title: A novel pathway for alternative splicing: Identification of an RNA intermediate  
A:Reference number: I57509; MUID:90377216; PMID:2398891  
A:Accession: I57509  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-17 <RES>

A;Cross-references: GB:M58688; NID:g202875; PIDN:AAA40726.1; PID:g554413  
C:Gene: AMPD1  
A:Introns: 8/1; 12/1  
C:Superfamily: AMP deaminase  
C:Keywords: hydrolase; muscle

Query Match 48.5%; Score 1505; DB 1; Length 747;  
Best Local Similarity 51.2%; Pred. No. 7.3e-104; Indels 16; Gaps 5;  
Matches 285; Conservative 87; Mismatches 169; Indels 16; Gaps 5;  
QY 19 PKPNPNPTTY--VPEPKSEHVFTQVDGVHVVADKDCTE-----SIYPVADATTTFFDLHY 72  
DB 188 PKKGEDPFRREDLPANLGVHL-KMKGGVYIYPDEAAASRDEPKPYENLDDFLDDMNF 246  
QY 73 ILRVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAHPDRDFNVNKKVDTVHH 132  
DB 247 LLALIAAGPVKTYTHRRLLKFLSSKFQVHQLNEMDELKELKNPNPHRDFNCRKVDTHHA 306  
QY 133 SACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVFESLDLTGYDLNVLDLVDHAK 192  
DB 307 AACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVFESLDLTGYDLNVLDLVDHAK 366  
QY 193 STPHRFDFENLKNPCQSGRLREIFLQDNLIQGRFLAELTKQVFSDLASASKYQMAEYRI 252  
DB 367 QTFQRFDFENLKNPCQSGRLREIFLQDNLIQGRFLAELTKQVFSDLASASKYQMAEYRI 426  
QY 253 SIYGRKQSEWDQLASWVNNELHSGNVMLVQIPRLYNYKEMGIIVTSFQNLNDLNFVL 312  
DB 427 SIYGRSPDEMSKLSWVFNVRNRYCPNMTWMIQVRIYDVFRSKNFPFLPHFGKMLNIFLEV 486  
QY 313 FEVTIPASHPOLHVLKQVVGDLVDDESKPERRPTKGM-----PTPEQWNTNVP 367  
DB 487 FEATINPQTHPDLSVFLKHTGTGSDVDDSK-----HSGHMFSSKSPKPEWNTNVP 542  
QY 368 YIAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHDLAATFLCHNISHGINLRKSP 427  
DB 543 YIAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHDLAATFLCHNISHGINLRKSP 602  
QY 428 VLOYLYLYGOIGLAMSPLNSLFLDYHRNPFTFFQRLNVLSTDDPLQHLTKPLV 487  
DB 603 VLOYLYLYGOIGLAMSPLNSLFLDYHRNPFTFFQRLNVLSTDDPLQHLTKPLV 662  
QY 488 EYYSIAASLWKLSSCDLCEARNVYQSGFSHAKAHWIGKYNFKRGPAGNDTHRTNVP 547  
DB 663 EYYSIAASLWKLSSCDLCEARNVYQSGFSHAKAHWIGKYNFKRGPAGNDTHRTNVP 722  
QY 548 IRVQFREMIRNEMKLV 564  
DB 723 IRMAYRYETWCYELNLI 739

RESULT 10  
I39444  
AMP deaminase (EC 3.5.4.6) - human  
N;Alternate names: myoadenylate deaminase  
C:Species: Homo sapiens (man)  
C:Date: 21-Feb-1997 #sequence\_revision 21-Feb-1997 #text\_change 22-Jun-1999  
C:Accession: I39444  
R;Sabina, R.L.; Fishbein, W.N.; Pezeshkpour, G.; Clarke, P.R.; Holmes, E.W.  
Neurology 42, 170-179, 1992  
A;Title: Molecular analysis of the myoadenylate deaminase deficiencies.  
A;Reference number: I39444; MUID:92131279; PMID:1370861  
A;Accession: I39444  
A;Status: preliminary; translated from GB/EMBL/DBU  
A;Molecule type: mRNA  
A;Residues: 1-747 <RES>  
A;Cross-references: GB:M60092; NID:g178543; PIDN:AAA57281.1; PID:g178544  
C:Gene: AMPD1  
A;Cross-references: GDB:113677; OMIM:102770  
A;Map position: lp13-lp13  
C:Superfamily: AMP deaminase

C:Keywords: hydrolase

Query Match 48.3%; Score 1497; DB 2; Length 747;  
Best Local Similarity 51.4%; Pred. No. 2.9e-103; Indels 22; Gaps 8;  
Matches 289; Conservative 89; Mismatches 162; Indels 22; Gaps 8;  
QY 18 TP--KPENPFF--TVPEPKSEHVFTQVDGVHVVY-----ADKDCATESI-YPVADATTTFF 67  
DB 185 TTPVKKGEDPFRDNLNLPENLGVHL-KMKDGVVYVYVNEAAVSKDEPKLPYENLDD--TFL 241  
QY 68 TDLHYILRVTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAHPDRDFNVNKKV 127  
DB 242 DDMNELLALIAAGPVKTYTHRRLLKFLSSKFQVHQLNEMDELKELKNPNPHRDFNCRKVD 301  
QY 128 THVHSAACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVFESLDLTGYDLNVLDL 187  
DB 302 THVHSAACMNOKHLRLFTKSLRKEPDEWIFRDGTVMYTLKEVFESLDLTGYDLNVLDL 361  
QY 188 VHAKSTPHRFDFENLKNPCQSGRLREIFLQDNLIQGRFLAELTKQVFSDLASASKYQ 247  
DB 362 VHAGRQTFQRFDFENLKNPCQSGRLREIFLQDNLIQGRFLAELTKQVFSDLASASKYQ 421  
QY 248 AEYRISIVGRKQSEWDQLASWVNNELHSGNVMLVQIPRLYNYKEMGIIVTSFQNLNDL 307  
DB 422 AEPLRSIYGRSPDEMSKLSWVFNVRNRYCPNMTWMIQVRIYDVFRSKNFPFLPHFGKMLN 481  
QY 308 IFVLPFVETIDPASHPOLHVLKQVVGDLVDDESKPERRPTKGM-----PTPEQWNTNVP 362  
DB 482 IFMPVFEATINPQTHPDLSVFLKHTGTGSDVDDSK-----HSGHMFSSKSPKPEWNTNVP 537  
QY 363 NPATSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHDLAATFLCHNISHGIN 422  
DB 538 NPATSYAYCYANLFTLNKLRSGKMTTIKFRPHAGAGDVHDLAATFLCHNISHGIN 597  
QY 423 LRKSVLYLYGOIGLAMSPLNSLFLDYHRNPFTFFQRLNVLSTDDPLQHLTKPLV 482  
DB 598 LRKSVLYLYGOIGLAMSPLNSLFLDYHRNPFTFFQRLNVLSTDDPLQHLTKPLV 657  
QY 483 KEPLVEEYIAAQQVFKLSTCDMCEARNVYQSGFSHAKAHWIGKYNFKRGPAGNDTHRTNVP 542  
DB 658 KEPLVEEYIAAQQVFKLSTCDMCEARNVYQSGFSHAKAHWIGKYNFKRGPAGNDTHRTNVP 717  
QY 543 TNPVHIRVQFREMIRNEMKLV 564  
DB 718 TNPVHIRVQFREMIRNEMKLV 739

RESULT 11  
S59996  
AMP deaminase (EC 3.5.4.6) 2 isoform L - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 22-Jun-1999  
C:Accession: S59996  
R;van den Bergh, F.; Sabina, R.L.  
Biochem. J. 312, 401-410, 1995  
A;Title: Characterization of human AMP deaminase 2 (AMPD2) gene expression reveals alter  
A;Reference number: S59996; MUID:96103174; PMID:8526848  
A;Accession: S59996  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-605 <RNA>  
A;Cross-references: EMBL:U16269; NID:g608496; PIDN:AAB06511.1; PID:g608497  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994  
C:Genetics:  
A;Introns: 175/1; 222/2; 360/3; 404/3  
C:Superfamily: AMP deaminase  
C:Keywords: alternative splicing; hydrolase

Query Match 38.3%; Score 1186.5; DB 2; Length 605;  
Best Local Similarity 50.6%; Pred. No. 2.6e-80; Indels 19; Gaps 4;  
Matches 222; Conservative 67; Mismatches 131; Indels 19; Gaps 4;  
QY 1 PRVRVAPWEKEVINDEPCTPKPNPNFTYVPEPKSEHVFTQVDGVHVVY-----DKDCATES 56

```

Db      181 PALEQHPYEH-----CESTWPGDGLG-----GLRWVGVVHYTRRDEHCSV 226
QY      57 IYPVADATTFDLDHLYLRTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAP 116
Db      227 ELPYDLOQFVADVNVLMALIINGPIKSCYRRLOQLVSSKFQMHVLLNEMKELAAQKVP 286
QY      117 HRDFYNRKVDTHVHSHACWQKHLRLFTKSKLRKEPDDEVIVFRDGTMTLKEVPSLDL 176
Db      287 HRDFYNRKVDTHHSHACWQKHLRLFTKSKLRKEPDDEVIVFRDGTMTLKEVPSLDL 346
QY      177 TGVDLNVLDDVHADKSTHRRFDKFNLYNCPGQSRLEIFLKQDNLIOGRFLAELTKOV 236
Db      347 TAYDLSVDLVDHADKSTHRRFDKFNLYNCPGQSRLEIFLKQDNLIOGRFLAELTKOV 406
QY      237 FSDLSASKYQMAEYRISIIYGRKQSEWDQLASWIVNNELHSGNVVLMVQIPRLYNNYKENG 296
Db      407 MSDLBESKYQNAELRLSIYGRSRDEWDKLARWAVHVRHSPNVRLVQVPRLFVYRTKG 466
QY      297 IVTSFQNLNDNFVPLFEVITIDPASHPQLHVLKQVVGDLVDVDDSKPERRP-TKHMPTP 355
Db      467 QLANQOEMLENIFLPLFEATVHPASHPELHLEFLEHVDGFSVDVDDSKPENHVFNLESPLP 526
QY      356 EOWTVNVPNPAFYAYCYANLFTLNKLRSGMTTIKPERHAGAGDVVDHLAATFLLCH 415
Db      527 EAWVEDNPPAYLYLTFFANWMLNHLRRQGRGFTFVLRRPHCGAGPIHHLVSAFMLAE 586
QY      416 NISHGINLRKSPVLOLYYY 434
Db      587 NISHGULLRKAPVLOLYYY 605

RESULT 12
S50801
AMP deaminase homolog YJL070c - yeast (Saccharomyces cerevisiae)
N:Alternate names: probable membrane protein YJL070c; protein HRD888; protein J1095
C:Species: Saccharomyces cerevisiae
C>Date: 23-Aug-1995 #sequence revision 08-Sep-1995 #text_change 19-Apr-2002
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
Yeast 11, 57-60, 1995
A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevisiae
A:Reference number: S50799; MUID:95282514; PMID:7762302
A:Accession: S50801
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-888 <VAN>
A:Cross-references: EMBL:Z34288; NID:g498992; PIDN:CAA84052.1; PID:g498996
R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
submitted to the EMBL Data Library, June 1994
A:Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Saccharo
A:Reference number: S47117
A:Accession: S47120
A:Molecule type: DNA
A:Residues: 1-888 <VAN>
A:Cross-references: EMBL:Z34288; NID:g498992; PIDN:CAA84052.1; PID:g498996
R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.
submitted to the Protein Sequence Database, September 1995
A:Reference number: S56835
A:Accession: S56844
A:Molecule type: DNA
A:Residues: 1-888 <POH>
A:Cross-references: EMBL:Z49345; NID:g1008222; PIDN:CAA89362.1; PID:g1008223; MIPS:YJL07
R:Sor, F.J.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57731
A:Accession: S57733
A:Molecule type: DNA
A:Residues: 1-888 <SOR>
A:Cross-references: EMBL:X88851; NID:g895892; PIDN:CAA61309.1; PID:g895895
C:Genetics:
A:Cross-references: SGD:S0003606

```

```

A:Map position: 10L
C:Keywords: transmembrane protein

Query Match      20.2%; Score 627; DB 2; Length 888;
Best Local Similarity 30.1%; Pred. No. 1.9e-38;
Matches 184; Conservative 94; Mismatches 207; Indels 126; Gaps 21;

QY      66 FFTDLHYLRTAAGNTRTVCHNRLNLEHKKFHLMLNADREFLAQKTAPHRDYPYNRK 125
Db      281 FRDQFAYIIELTQSHKFNVEGRKLSYLLDKRFELFOYLNKSKKELILANKVVPYRDFYNSRK 340
QY      126 VDTTHVHSHACWQKHLRLFTKSKLRKEPDDEVIVFRDGTMTLKEVPSLDL 171
Db      341 VDRDLSLGGCISQRLSEYIWEKINLEPERIV-YQDPETSKRLSLRDIQFQCSSNDQPI 399
QY      172 -----ESLDLTGYDLNVDLVDHADKSTHRRFDKFNLYNCPGQSRLEIFLKQDNL 222
Db      400 AIGLKLIDDELD---WVRNIYLDIDYHLTPNKVAKLVGKEMRF-----YLLAKVLEFDN 451
QY      223 LIQGRFLAEL-TKQVFSDSLASKYQMAEYRIS--IYGRKQSEWDQLASWIVNNELHSGNV 279
Db      452 FIEGEYLAELFIKVIHILEKSKYQLAQVSNVQFYSSGEDWYKKFSQWLLRWKLVSYNI 511
QY      280 VMLVQIPRLY-NVYKEMGIIVTSFQNLNDNFVPLF-----EVTIDPASHPO---LHVFLK 330
Db      512 RWNIIQIARIFPKLFKE-NVVSNFQEFDLIFENPLFTLEKEQLPIDSSVNTDIIGLOFFLS 570
QY      331 QVVGDLVDVDDSKPBE--RRPTKHMPTPEQWT-NVFNPAFSYVAYCYANLFTLNKLRSG 387
Db      571 NVCSMDLVKESDEYVWKEFTDMNCKPFWTAQGDNPVTAHMYIYKSLAKVNLRSQN 630
QY      388 GMTTIKFRPHAGE-----AGDVHDLAATFLLCHNISHGINLRKSP----- 427
Db      631 LQNTITLRNYCPSLSSRTSQFGVDLYFTDQVESLVCNLLC-----NGLLQVEPLWDRTAT 686
QY      428 VLQILYVYLGQGLAWSPLSNNSL-----PL----- 452
Db      687 MIQVLYFLQIPILAAPLSSVSLNSQSKSTFLKNKNVLLHEDYLDKQDQETAKINPSRDITV 746
QY      453 ---DYHNRNPPPTFFQRLGNVLSLTDQPL-QIHLTKEPLVEEYSYTAASLWKLSSCDLCEI 507
Db      747 GEQSYETNPWPKMFKMGLKLSLSKSLYNSVYTLPELIBEYSVAASIIYLLNPTDLCCL 806
QY      508 ARNSVYQSGFSGHALKAHWIGKN-----YFKGKPGAGND-----IHRTNVPHIRVQFR 553
Db      807 SRTSVLSSGYEGWYKAHWIGVGVKKAPYFEEVNGVDNWDYDTAKDTSIKHNVPIMRRYR 866
QY      554 EMIWRNEMKLV 564
Db      867 KETLDQEWNVFV 877

RESULT 13
S44546
probable membrane protein YBR284w - yeast (Saccharomyces cerevisiae)
N:Alternate names: hypothetical protein YBR2021
C:Species: Saccharomyces cerevisiae
C>Date: 08-Jun-1994 #sequence revision 09-Sep-1994 #text_change 19-Apr-2002
R:Holmstrom, K.; Brandt, T.; Kallese, T.
Yeast 10(Suppl.A), S47-S62, 1994
A:Title: The sequence of a 32420 bp segment located on the right arm of chromosome II fr
A:Reference number: S44537; MUID:94378722; PMID:8091861
A:Accession: S44546
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-797 <HOL>
A:Cross-references: EMBL:X76053; NID:g600025; PID:g429129
R:Brandt, T.; Christiansen, C.; Holmstrom, K.; Kallese, T.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S46157
A:Accession: S46166
A:Molecule type: DNA

```

A;Residues: 1-797 <BRA>  
A;Cross-references: EMBL:Z36153; NID:G536731; PID:G536732; MIPS:YER284w  
C;Genetics:  
A;Cross-references: SGD:S0000488  
A;Map position: 2R  
C;Superfamily: AMP deaminase  
C;Keywords: transmembrane protein  
F;501-517/Domain: transmembrane #status predicted <TM1>  
F;627-643/Domain: transmembrane #status predicted <TM2>

Query Match 20.2%; Score 625; DB 2; Length 797;  
Best Local Similarity 30.6%; Pred. No. 2.3e-38;  
Matches 182; Conservative 86; Mismatches 213; Indels 114; Gaps 21;

QY 66 FFDLHVLIRVTAAGNTRIVCHNRLNLEHKKFHLNADREFLAQKTAHPDRDFVNVK 125  
DB 219 FREDFWCLUIRDSLSRFSEKRLQVLNKFVFOHLHSKEMROSKVPHKDFVNCRK 278  
QY 126 VDPHVHSACMOKHLLRIKSLRKEPDEV--IFRDGTMYTLKEVFE-SLDLTG--- 178  
DB 279 IDLNLGSCFSQWQLTEFIWTKLKEPDRVIHQAF-NGSHITLSQLFKVNFEEGTQFFN 337  
QY 179 -----YDLNVLDLVHADKTFHRDPKFNKYNPCGQSRLEIFL 218  
DB 338 GLKIIDSFLEWYKVYLAHYLHVNDMEIHTGS-----HGKQLRY-----YLIAKTEL 386  
QY 219 KODNLQGRFLAELTKQ-VFSDLSASKYQMAEYRISI-----YGRKQSE--WDQLASWIV 270  
DB 387 EFDNYINGEYLAELLTKFLIKPOESKYQLCQSVDFQFYLVHVDNVDNMMVFNWLN 446  
QY 271 NNELHSGNVWLVQIPRLXNVYKEMGIVTSFQNLNLIIVPLF--EVTIDPASHPOLHVF 328  
DB 447 HYNIPSNIRWNRIRISRIPELVHTCKVKNFQVYLNLIKPLFNAENYHLKSLGPIILKF 506  
QY 329 LKQVGLDLV--DDESKPERRPKHMPTEQWNT-VFNPAFSYVAYVCYANLFTLNKLR 385  
DB 507 LQSVSSIDLQIDTNYIWKNFYAVSCLPKDWTSGGDNPTISQYMYVYVNLTKLNHIQ 566  
QY 386 -----SKGTTIKFRPHAGEAGDVHDLAATFLI-CHNISHGINL 423  
DB 567 ALHQTFTLRSSCSPTSMNRTSOFNTLNPTEH-----TEALNNFLACGGFLNAENL 620  
QY 424 -RKSPVLQVLYLQIGLAMSPLSN-----NSFLD----YHRNPFTFFQ 464  
DB 621 WNAPPSLVLYLYLSQIPMVVAPLNSIVDSKPTMLQEQAPGLVLEPSKPKYKKNPFMKFFE 680  
QY 465 RGLNVSLSTDDPLQIHL-LKEPLVEEYSTAASLWKLSSCDLCEIARNSVYQSGFHALKA 523  
DB 681 MGFKISLSSESILYNNSTKEPIIEEYSVAASIYRLHSADLCLELNSVITSGFSTLKN 740  
QY 524 HWIG-----KNYFKRGPA-----GNDIHRNTVPHIRVOPREMIWRNEMKLV 564  
DB 741 KWLGVSLASHDYEVNTGFDVKWYDCKPNTSLHNHVPPIRRQYRSSTLAGEWRLLI 795

## RESULT 14

A37056  
AMP deaminase (EC 3.5.4.6), brain - rat (fragment)  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 25-Jan-1991 #sequence\_revision 25-Jan-1991 #text\_change 22-Jun-1999  
C;Accession: A37056  
R;Morisaki, T.; Sabina, R.L.; Holmes, E.W.  
J. Biol. Chem. 265, 11482-11486, 1990  
A;Title: Adenylate deaminase. A multigene family in humans and rats.  
A;Reference number: A37056; MUID:90307656; PMID:2365682  
A;Accession: A37056  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-88 <MOR>  
A;Cross-references: GB:M38126; NID:G202880; PID:AAA40728.1; PID:G202881  
C;Superfamily: AMP deaminase  
C;Keywords: hydrolase

Query Match 11.3%; Score 350; DB 2; Length 88;  
Best Local Similarity 72.7%; Pred. No. 2.8e-19;  
Matches 64; Conservative 7; Mismatches 17; Indels 0; Gaps 0;

QY 396 PHAGEAGDVHDLAATFLCHNISHGINLRKSPVLQVLYLQIGLAMSPLSNLSFLDYH 455  
DB 1 PHCGEAGPIHHLVSAFMAENISHGLLRKAPVLYLYLAQIGIAMSPLSNLSFLSYH 60  
QY 456 RNPPTFFQRLNVSLSTDDPLQIHLTK 483  
DB 61 RNPPEYLSRGLMVSLSTDDPLQPFHTK 88

RESULT 15  
G97269  
adenosine deaminase [imported] - Clostridium acetobutylicum  
C;Species: Clostridium acetobutylicum  
C;Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 30-Sep-2001  
C;Accession: G97269  
R;Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl.  
A;Reference number: A96900; MUID:21359325; PMID:21359325  
A;Accession: G97269  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-334 <KUR>  
A;Cross-references: GB:AE001437; PIDN:AAK80946.1; PID:gl5026062; GSPDB:GN00168  
A;Experimental source: Clostridium acetobutylicum ATCC824  
C;Genetics:  
A;Gene: CAC3005  
C;Superfamily: adenosine deaminase

Query Match 4.8%; Score 149; DB 2; Length 334;  
Best Local Similarity 27.3%; Pred. No. 0.0017;  
Matches 65; Conservative 39; Mismatches 106; Indels 28; Gaps 11;

QY 314 EVTIDPASHPOLHVLKQV--GLD-LVDDSKPERRPT-----KHMPTEQWNTVNP 364  
DB 100 EIRFAPFQHTQDNLNENDVVEAALQDGESKLGHSNLLILCSLRHDPV-ERSIDLNL 158  
QY 365 AFSYVAYVCYANL-----FTLNKLRESKGM---TTIKFRPHAGEAGDVHDLAATFLCH 415  
DB 159 ANSYNEGVCAYDLNAGNESDFPELHKEAFDAYDNGIKITIHAGETGAENILKSIKLH 218  
QY 416 --NISHGINLRKS-PVLQVLYLQIGLAMSPLSN--NSFLDYHRNPFTFFQRLNV 470  
DB 219 ADRIHGIFAYKSEILOYVIE-NQVPLEMCPKSNVDTKAVKNYKHPFKKYFDLGKVT 277  
QY 471 LSTDDPLQIHLTKPELVEEYSTAASLWKLSSCDLCEIARNSVYQSGFHALKAHWIG 528  
DB 278 LNTDRT---VSNVSLVDEYLNLANIFDGTIEIKTVIRNGISASFATEEFKVNLLKK 332

Search completed: December 15, 2003, 13:28:40

Job time : 23 secs